

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:54:17 ; Search time 304 Seconds

(without alignments)
1340.366 Million cell updates/sec

Title: US-09-881-736-2

Perfect score: 3243

Sequence: 1 MDTMMLNVRNLFFQQLVRRVE... SKSANTLGRQGNEFFASPMLK 632

Scoring table: BL0SUM62

Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending Patents AA Main:
1: /rgn2_6/ptodata/1/paa/ructus_comb.pep:*

2: /rgn2_6/ptodata/1/paa/us06_comb.pep:*

3: /rgn2_6/ptodata/1/paa/us07_comb.pep:*

4: /rgn2_5/ptodata/1/paa/us080_comb.pep:*

5: /rgn2_6/ptodata/1/paa/us081_comb.pep:*

6: /rgn2_6/ptodata/1/paa/us082_comb.pep:*

7: /rgn2_6/ptodata/1/paa/us083_comb.pep:*

8: /rgn2_6/ptodata/1/paa/us084_comb.pep:*

9: /rgn2_6/ptodata/1/paa/us085_comb.pep:*

10: /rgn2_6/ptodata/1/paa/us086_comb.pep:*

11: /rgn2_6/ptodata/1/paa/us087_comb.pep:*

12: /rgn2_6/ptodata/1/paa/us088_comb.pep:*

13: /rgn2_6/ptodata/1/paa/us089_comb.pep:*

14: /rgn2_6/ptodata/1/paa/us090_comb.pep:*

15: /rgn2_6/ptodata/1/paa/us091_comb.pep:*

16: /rgn2_6/ptodata/1/paa/us092_comb.pep:*

17: /rgn2_6/ptodata/1/paa/us093_comb.pep:*

18: /rgn2_6/ptodata/1/paa/us094_comb.pep:*

19: /rgn2_6/ptodata/1/paa/us095_comb.pep:*

20: /rgn2_6/ptodata/1/paa/us096_comb.pep:*

21: /rgn2_6/ptodata/1/paa/us097_comb.pep:*

22: /rgn2_6/ptodata/1/paa/us098_comb.pep:*

23: /rgn2_6/ptodata/1/paa/us099_comb.pep:*

24: /rgn2_6/ptodata/1/paa/us100_comb.pep:*

25: /rgn2_6/ptodata/1/paa/us101_comb.pep:*

26: /rgn2_6/ptodata/1/paa/us102_comb.pep:*

27: /rgn2_6/ptodata/1/paa/us60_comb.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3243	100.0	632	21 US-09-791-537-1796 Sequence 1796, Ap
2	3243	100.0	632	22 US-09-881-736-2 Sequence 2, Appli
3	3237	99.8	632	20 US-09-661-377-585 Sequence 585, App
4	3233	99.7	632	18 US-09-488-755A-3192 Sequence 3192, Ap
5	3230	99.6	632	21 US-09-991-537-72082 Sequence 72082, A
6	3230	99.6	632	22 US-09-833-790-413 Sequence 413, App

ALIGNMENTS

RESULT 1

US-09-791-537-1796

; Sequence 1796, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Bionomix, Inc.

; APPLICANT: Dake, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-07-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 1796

; LENGTH: 632

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-791-537-1796

; QUERY

; MATCH

; LENGTH

; DB

; QY

; Matches

;保守性

; 0;

; Mismatches

; 0;

; Indels

; 0;

; Gaps

; 0;

; Sequence 427, App

; Sequence 36304, A

; Sequence 6751, Ap

; Sequence 14587, A

; Sequence 1045, Ap

; Sequence 30960, A

; Sequence 49124, A

; Sequence 30526, A

; Sequence 967, Ap

; Sequence 968, App

; Sequence 294, App

; Sequence 699, App

; Sequence 9071, Ap

; Sequence 2570, Ap

; Sequence 4516, A

; Sequence 4589, A

; Sequence 50263, A

; Sequence 44317, A

; Sequence 6764, Ap

; Sequence 3866, Ap

; Sequence 6453, Ap

; Sequence 76252, A

; Sequence 6, Appli

; Sequence 3547, Ap

; Sequence 35867, A

; Sequence 77334, A

; Sequence 3493, Ap

; Sequence 65301, A

; Sequence 4632, Ap

; Sequence 4551, Ap

; Sequence 150251, Ap

; Sequence 901, App

; Sequence 598, App

Query Match 99.6%; Score 3230; DB 21; Length 632;
 Best Local Similarity 99.7%; Pred. No. 3..5e-262; Length 632;
 Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDTMMLNVLNLFQDQVRRVEILSBCNEVOFTIQLAKDFEPRKKWORTDHELGKVKDLMK 60
 1 MDTMMLNVLNLFQDQVRRVEILSBCNEVOFTIQLAKDFEPRKKWORTDHELGKVKDLMK 60

Db 61 AETERSALDWKLKHARNQDVETRKQRREADCEKLERIOQLIREMKTSGSIOLSE 120
 61 AETERSALDWKLKHARNQDVETRKQRREADCEKLERIOQLIREMKTSGSIOLSE 120

Db 61 AETERSALDWKLKHARNQDVETRKQRREADCEKLERIOQLIREMKTSGSIOLSE 120
 61 AETERSALDWKLKHARNQDVETRKQRREADCEKLERIOQLIREMKTSGSIOLSE 120

QY 121 QKSALAFLNRGQPPSSNAGNKRSLTIDESGSILSDISFDKTDSDLWDSSLVKTFKLKR 180
 121 QKSALAFLNRGQPPSSNAGNKRSLTIDESGSILSHISFDKTDSDLWDSSLVKTFKLKR 180

Db 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240
 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240

Db 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240
 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240

QY 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300
 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300

Db 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300
 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300

QY 301 VPGCKRIKGKLSKCRDRVSHEPCRCPUCPICTLGTPKIGEMLADVSQSP 360
 301 VPGCKRIKGKLSKCRDRVSHEPCRCPUCPICTLGTPKIGEMLADVSQSP 360

Db 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420
 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420

Db 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420
 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420

QY 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480
 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480

Db 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480
 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480

QY 481 QRVASOPHTMDVANLAKYFGPTVVAHAVPNPDVTMSDIKROPKVVERLISPLEWS 540
 481 QRVASOPHTMDVANLAKYFGPTVVAHAVPNPDVTMSDIKROPKVVERLISPLEWS 540

Db 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600
 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600

Db 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600
 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600

QY 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632
 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632

Db 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632
 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632

RESULT 6
 US-09-833-790-413
 ; Sequence 413, Application US/09833790
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Lodges, Michael J.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Sechrist, Heather
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Indrias, Carol Y.
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121..512
 ; CURRENT APPLICATION NUMBER: US/09/833,790
 ; CURRENT FILING DATE: 2001-04-11
 ; NUMBER OF SEQ ID NOS: 440
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 413
 ; LENGTH: 632
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ;
 US-09-833-790-413

RESULT 7
 US-09-833-790-427
 ; Sequence 427, Application US/09833790
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Sechrist, Heather
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Indrias, Carol Y.
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121..512
 ; CURRENT APPLICATION NUMBER: US/09/833,790
 ; CURRENT FILING DATE: 2001-04-11
 ; NUMBER OF SEQ ID NOS: 440
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 427
 ; LENGTH: 570
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ;
 US-09-833-790-427

Query Match 99.6%; Score 3230; DB 22; Length 632;
 Best Local Similarity 99.7%; Pred. No. 3..5e-262; Length 632;
 Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 121 QKSALAFLNRGQPPSSNAGNKRSLTIDESGSILSDISFDKTDSDLWDSSLVKTFKLKR 180
 121 QKSALAFLNRGQPPSSNAGNKRSLTIDESGSILSHISFDKTDSDLWDSSLVKTFKLKR 180

Db 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240
 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240

Db 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240
 181 EKRRTSRQFDGPGPPVKTTSIGSAVODGNESIVAKTVTVNDGGPTEAVSTIEVP 240

QY 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300
 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300

Db 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300
 241 YWTRSRRKGTQPNNSDTLNSQLEPRTEDSGTPOSNGGHLHDVSKTVIKPESC 300

QY 301 VPGCKRIKGKLSKCRDRVSHEPCRCPUCPICTLGTPKIGEMLADVSQSP 360
 301 VPGCKRIKGKLSKCRDRVSHEPCRCPUCPICTLGTPKIGEMLADVSQSP 360

Db 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420
 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420

Db 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420
 361 MIPSVWHCUNEIQORGLETGLYRISGCDRTVKELEKFLRVKIVPLSKVDDIHATCS 420

QY 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480
 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480

Db 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480
 421 LKKDFLRNKEPLITFLRNRAFMEAETDEDSIAAMQAVGELPQANRDTLAFMLH 480

QY 481 QRVASOPHTMDVANLAKYFGPTVVAHAVPNPDVTMSDIKROPKVVERLISPLEWS 540
 481 QRVASOPHTMDVANLAKYFGPTVVAHAVPNPDVTMSDIKROPKVVERLISPLEWS 540

Db 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600
 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600

Db 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600
 541 QFMVMEQENIDPLHVTIENNAFSPTQTPKIVSLLGPVTPHEOLLKPPSSSLSQVRVS 600

QY 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632
 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632

Db 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632
 601 TLTKNTPRGSKSKSATNLGRQGFASPLK 632

Best Local Similarity 99.8%; Pred. No. 4.5e-235; Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0; Db 61 KAPERSALDVKLKHARNQDVIEKRRQRAEADCEKLERQIQLIREMLMDTSQSIQLOSEK 120

63 TERSALDKLKLKHARNQDVIEKRRQRAEADCEKLERQIQLIREMLMDTSQSIQLOSEK 122

1 TERSALDKLKLKHARNQDVIEKRRQRAEADCEKLERQIQLIREMLMDTSQSIQLOSEK 60

123 SALAFLNQRQPSSENAGNKRSLTIDESGSILSFDKTDLSLDSLURTEFKLKKREK 182.

61 SALAFLNQRQPSSENAGNKRSLTIDESGSILSFDKTDLSLDSLURTEFKLKKREK 120

183 RRSTSRQFVGPPGPKVURSIGSAVDDQESTVAKTTVWPNQGPPIEVSTIEVWPW 242

121 RRSTSRQFVGPPGPKVURSIGSAVDDQESTVAKTTVWPNQGPPIEVSTIEVWPW 180

181 TRSRKRTTLOPWNNSDTSLSRQLEPRTEDSVGTPOSNGMRLHDFVSKTVIKESCP 240

303 CGRKIKEKSLKRCRDVSHPECRDPLCIPMIGPVKLGMBMLADFSQSPMI 362

243 KDFRNKLKEPLTRLNRAFMEAETDENSIAAMYQAVGELPOANRDTLAFLMHLQR 482

361 CGRKIKEKSLKRCRDVSHPECRDPLCIPMIGPVKLGMBMLADFSQSPMI 300

363 PSTWVHCNEIEORGTLTGYRISGCDRUYKEKEFLKVTKVLLSKYDDIHACSL 422

301 PSTWVHCNEIEORGTLTGYRISGCDRUYKEKEFLKVTKVLLSKYDDIHACSL 360

423 KDFRNKLKEPLTRLNRAFMEAETDENSIAAMYQAVGELPOANRDTLAFLMHLQR 482

361 CGRKIKEKSLKRCRDVSHPECRDPLCIPMIGPVKLGMBMLADFSQSPMI 420

483 VASPHTRKDVANLAKVFGPTIVAHAVNPDPVMSQDIKROPKVERLSSPLEWQF 542

421 VASPHTRKDVANLAKVFGPTIVAHAVNPDPVMSQDIKROPKVERLSSPLEWQF 480

543 MMYEQENIDPLHYTENSIAFSTPOTPDIVSLLGVTTPHOLKTPSSSLSQRSTL 602

481 MMVQENIDPLHYTENSIAFSTPOTPDIVSLLGVTTPHOLKTPSSSLSQRSTL 60

603 TKNPREFSKSKSATNLGQNGFFASPMLK 632

541 TKNPREFSKSKSATNLGQNGFFASPMLK 570

RESULT 8

-09-791-537-36304

Sequence 36304, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Bionomix, Inc.

APPLICANT: Bebe, Berek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER

FILE REFERENCE: 264/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 36304

LENGTH: 628

TYPE: PRT

ORGANISM: Mus musculus

Query Match Score 84.3%; Score 2734; DB 21; Length 628;

Best Local Similarity 84.4%; Pred. No. 2e-220; Mismatches 47; Mismatches 46; Indels 6; Gaps 3; Matches 534; Conservative 47; Mismatches 46; Indels 6; Gaps 3; SEQ ID NO 4

1 MDPMMLNVRNLFEOLYRVEILSEGNE-YQFIQLAKDFEDFRKKWQRTDHGLGVKDLIM 59

1 MDPMMLNVRNLFEOLYRVEILSEGNE-YQFIQLAKDFEDFRKKWQRTDHGLGVKDLIM 59

60 KAETERSALDKLKHARNQDVIEKRRQRAEADCEKLERQIQLIREMLMDTSQSIQLOSE 119

RESULT 9

US-09-881-736-4

; Sequence 4, Application US/09881736

; GENERAL INFORMATION:

; APPLICANT: Glotzen, Michael

; APPLICANT: Jantsch-Plunger, Verena

; APPLICANT: Romano, Alper

; APPLICANT: Mishima, Maanori

; APPLICANT: Kaitina, Susanne

; TITLE OF INVENTION: CYK-4 polypeptides, DNA molecules encoding them and their use

; FILE REFERENCE: 0652_226001/ERS-AES

; CURRENT APPLICATION NUMBER: US/09/881,736

; CURRENT FILING DATE: 2001-06-18

; PRIOR APPLICATION NUMBER: EP 00 112 880.0

; PRIOR FILING DATE: 2000-06-19

; PRIOR APPLICATION NUMBER: EP 01 110 554.1

; PRIOR FILING DATE: 2001-04-30

; PRIOR APPLICATION NUMBER: 60/241,231

; PRIOR FILING DATE: 2000-10-18

; PRIOR APPLICATION NUMBER: To be determined

; PRIOR FILING DATE: 2001-06-13

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

US-09-881-736-4

; ORGANISM: Mus musculus

Query Match Score 84.3%; Score 2734; DB 22; Length 628;

Best Local Similarity 84.4%; Pred. No. 2e-220; Mismatches 47; Mismatches 46; Indels 6; Gaps 3; Matches 534; Conservative 47; Mismatches 46; Indels 6; Gaps 3;

Query Match: 34.0%; Score: 1103.5; DB: 26; Length: 243;
 Best Local Similarity: 97.4%; Pred. No.: 8; $P_{\text{e}} = 84$;

Search completed: July 3, 2003, 10:02:52
Job time : 307 secs

QY 1 MDTMMLNVLNEFQURVREILSEGNEVFIQLAKKDFEDERKKWRTDHEIGKVDLMK 60
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 ||| 4 MDTMMLNVLNLFEQURVREILSEGNEVFIQLAKKDFEDERKKWRTDHEIGKVDLMK 63
 Db 61 AETERSALDVKLKRNQDVETKRRQRABEADCKLERQTLIREMLCDTSQIOLSEE 120
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 ||| 64 AETERSALDVKLKRNQDVETKRRQRABEADCKLERQTLIREMLCDTSQIOLSEE 123
 QY 121 QKSALAFLNQGPSSNAGNKRSLTIDESGSIISDISFRKTDESIDWDSLSVKFLKKR 180
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 ||| 124 QKSALAFLNQGPSSNAGNKRSLTIDESGSIISDISFRKTDESIDWDSLSVKFLKKR 183
 Db 181 EKRSTSREFDGGPGPVKTRSGSAVGQNESIVAKTIVTVNGG 228
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 ||| 184 EKRSTSREFDGGPGPVKTRSGSAVGQNESIVAKTIVTVNGG 230

RESULT 15
PCP-US01-08656-9072

Sequence 9072, Application PC/US0108656

GENERAL INFORMATION:
APPLICANT: HYSEC, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-066
CURRENT APPLICATION NUMBER: PCP-US01/08656
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 09/522,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-11-26
NUMBER OF SEQ ID NOS: 10994
SOFTWARE: Custom
SEQ ID NO: 9072
LENGTH: 256
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (7)-(33)
OTHER INFORMATION: PROTEIN GTPASE DOMAIN ACTIVATION domain identified by OTHER INFORMATION: eMATRIX, accession number PD00930A, p-value=7.324e-15, rCT-US01-08656-9072

Query Match 30.0%; Score 971.5; DB 1; Length 256;
Best Local Similarity 84.0%; Pred. No: 1.2e-72;
Matches 205; **Conservative** 3; **Mismatches** 27; **Indels** 9; **Gaps** 4;

QY 357 QTSPNIPSPITVHCVCNEIEORGLTGEGLYLISCGCDRTVKEKLETKRVTKVPLISKVDDH 416
Db 1 QTSPNIPSPITVHCVCNEIEORGLTGEGLYLISCGCDRTVKEKLETKRVTKVPLISKVDDH 60

QY 417 AICSLKLDLFLRNKLEPLUTPLRNGAFAMEAETEDNSIAAMOAVGELPQANRDTLRL 476
Db 61 AICSLKLDLFLRNKLEPLUTPLRNGAFAMEAETEDNSIAAMOAVGELPQANRDTLRL 120

QY 477 MIHQVRQASPHTKNDVANLAKYFGPTIVAHAVENPDPTMSQDQIKR_OPKVYVERLLSTP 535
Db 121 MIHQVRQASPHTKNDVANLAKYFGPTIVAHAVENPDPTMSQDQIKR_OPKVYVERLLSTP 535

QY 536 L-EYKWSQFMNVEQI---NIDPLRVTEENSVAFSPQTQDVK---SLGPVTPHEQHLK 587
Db 181 LWEYWEFSFMDGWEGGTFDPLHVLENSNAFSTPOTPDIKAVPGGGLCVHFTAGEAEIQL 240

QY . 588 TPSS 591
Db 241 GPPS 244

GenCore version 5.1.6
copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Title: US-09-881-736-2
Perfect score: 3243

Sequence: 1 MDTMMMLNVRNLFEQLVRRVE. SKSATNLGRQGNFFASPMILK 632
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 20647115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%, Maximum Match 100%
Listing first 45 summaries

Database :	SPREMBL_21:*
1:	sp_archea:*
2:	sp_bacteria:*
3:	sp_fungi:*
4:	sp_human:*
5:	sp_invertebrate:*
6:	sp_mammal:*
7:	sp_mhc:*
8:	sp_organelle:*
9:	sp_phage:*
10:	sp_plant:*
11:	sp_rabbit:*
12:	sp_virus:*
13:	sp_vertebrate:*
14:	sp_unclassified:*
15:	sp_rvirus:*
16:	sp_bacteriapl:*
17:	sp_archeapl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3243	100.0	632	4 Q9P2W2
2	3237	99.6	632	4 Q9H0H5
3	3230	99.6	632	4 Q9WNW2
4	2919	90.0	570	4 Q9P250
5	2781	85.8	628	4 Q9BZ74
6	2734	84.3	628	4 Q9WVM1
7	1182	36.4	255	4 Q9H9L9
8	864	26.6	625	5 Q9V6U7
9	862	26.6	625	5 Q9N9Z9
10	547.5	16.9	681	5 Q9XUS9
11	431	13.3	354	5 Q95VRL
12	328	10.1	334	11 Q9IV57
13	322	9.9	459	4 Q96FB0
14	320	9.9	299	13 Q92153
15	308	9.5	695	3 Q94466
16	305.5	9.4	1388	5 Q8ST20

ALIGNMENTS

RESULT 1

Q9P2W2 PRELIMINARY; PRT; 632 AA.
ID Q9P2W2;
AC Q9P2W2;
DT 01-OCT-2000 (TRIMBLEL 15, last sequence update)
DT 01-OCT-2000 (TRIMBLEL 15, last annotation update)
DT 01-JUN-2002 (TRIMBLEL 21, last annotation update)
DE GTPase activating Protein.
GN MGCRAGAP.
OS Homo sapiens (Human).
RA Kawashima T., Hirose K., Satoh T., Kaneko A., Ikeda Y., Kaziro Y.,
Nosaka T., Kitamura T.;
RT "MGCRAGAP is involved in the control of growth and differentiation of
hematopoietic cells";
RT Blood 96:2116-2124(2000).
DR EMBL; AB030251; BAA0247.1; -.
DR HSSP; Q07960; IRGP.
DR InterPro; IPR002219; DAG_Pe-bind.
DR Intertpro; IPR000198; Rhocap.
DR Pfam; PF00130; DAG_Pe-bind; 1.
DR Pfam; PF00620; Rhogap; 1.
DR SMART; SM00109; Cl; 1.
DR SMART; SM00324; Rhogap; 1.
DR PROSITE; PS00479; DAG_Pe_bind_dom_1; UNKNOWN_1.
DR PROSITE; PS00081; DAG_Pe_bind_dom_2; 1.
DR SEQUENCE; 632 AA; 7100 MW; BC2175E5873B2B CRC64;

Query Match Best Local Similarity 100.0%; Score 3243; DB 4; Length 632; Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 1 MDTMMMLNVRNLFEQLVRRVEITSEGNEVQFOLAKUDFEDRKKWORTDHELGKYKDLMK 60 Q8ST20 caenorhabdi

1 MDTMMMLNTRNLFFOLYRVEILSEGNEVQFIOLAKDFEDFRKKWQDTHLGKYKDLMK 60
 DR Pfam; PF00630; RHOGAP; 1.
 DR SMART; SM01019; CL; 1.
 QY |||||||AETERSALDVYKLUHARNQDVETKRORRAEADCEKLERQIQLTREMIMCDTSGSIQSEE 120
 DR PROSITE; PS00479; DAG_Pe_BIND_DOM_1; UNKNOWN_1.
 DB 61 AETERSALDVYKLUHARNQDVETKRORRAEADCEKLERQIQLTREMIMCDTSGSIQSEE 120
 KW HYPOTHETICAL PROTEIN.
 QY |||||||OKSALAFLRNGQPSNSNAGNKLRTIDBGSILSDSFDTKTDLSMDSSLVKTFKLKR 180
 DR SQ SSEQUENCE 632 AA; 71026 MW; 032B7DF9C8A8F39D CRC64;
 DR 121 EKRSTSKQFDGPPGPVKTKRSIGSAVDQGNESIVAKTTVAVPNDGGPIEAUSTETVP 240
 DR 181 YWTRSRRTKGTLQPNNSDSTLNSRQLPRTEDSGVTPQSNQGMRLHDFFVSKVIREPSC 300
 DR 241 YWTRSRRTKGTLQPNNSDSTLNSRQLPRTEDSGVTPQSNQGMRLHDFFVSKVIREPSC 300
 DR 301 VPGKRRKIFGKUSLKCRDRCRVSHPECRDRCPCLPCIPPLIGTPVKIGEGMLADFSOTSP 360
 DR 361 MIPSIWVHCVNEEQRGITETGLYRISCGDRVKELEKFLRKVTPLSKVDDIACCS 420
 DR 361 MIPSIWVHCVNEEQRGITETGLYRISCGDRVKELEKFLRKVTPLSKVDDIACCS 420
 DR 421 LKKDFLRNLKEPLTFRNRAFMEAETTEDNSIAAMYQAVGELPOANRDTLAFLMHL 480
 DR 481 ORVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPPVTMSDQIKRQPKVERLLSLPLEWS 540
 DR 541 QFMVEQENIDPLVHVIENNSAATPQDPDIKVLIGLGPVTPHEQLKTPSSSSLQRVRS 600
 DR 541 QFMVEQENIDPLVHVIENNSAATPQDPDIKVLIGLGPVTPHEQLKTPSSSSLQRVRS 600
 DR 601 TLTKNTPRFGSKSKSATNLGROGNFFASPMLK 632
 DR 601 TLTKNTPRFGSKSKSATNLGROGNFFASPMLK 632

RESULT 2

09H0H5 PRELIMINARY; PRT; 632 AA.

ID 09H0H5 AC 09H0H5 DT 01-MAR-2001 (TREMBlrel. 15, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE Hypothetical 71.0 kDa protein.
 DR DKFPa34C01L.

N S Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN NCBI_TaxID=9606;
 RP [1]
 SEQUENCE FROM N.A.

TISSUE TESTIS;

MEDLINE="21154917"; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassnerhuber J., Glassl S.,
 RA Ansorge W., Boehler M., Blauecker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.W., Oettwaele B., Oettwaele B., Tampe J., Heubner D.,
 RA Wambets R., Korn B., Klein M., Pousta A.;
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 Analysis of 500 Novel Complete Protein Coding Human cDNAs,"
 RL Genome Res. 11:422-435 (2001);
 DR EMBL; All36794; CAB66728; -;
 DR HSSP; Q07960; IPRP;
 DR InterPro; IPR002219; DAG_Pe-bind.
 DR InterPro; IPR00198; RhogAP.
 DR Pram; PF00130; DAG_Pe-bind; 1.

RESULT 3

09NW2 PRELIMINARY; PRT; 632 AA.

ID 09NW2 AC 09NW2; DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DR CDNA FLJ20726 fis, clone HE13735
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 SEQUENCE FROM N.A.

RP Nakajima T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,

RA	Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.,	OS	Homo sapiens (Human).
RT	"NEDO human cDNA sequencing project.",	OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
RL	SUBMITTED (FEB-2000) to the EMBL/GenBank/DDBJ databases.	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DR	EMBL: AK00733; BAA91347.1;	OX	NCBI_TaxID=9606;
DR	HSSP: Q07960; IRG:	RN	[1]
DR	InterPro: IPR002219; DAG_PE-bind.	RP	SEQUENCE FROM N-A.
DR	Pfam: PF00130; DAG_PE-bind; 1.	RX	MEDLINE=20277482; PubMed=10819331;
DR	Pfam: PF00620; RhogAP; 1.	RA	Nazae T., Kikuno R., Ishikawa K., Hirosa M., Ohara O.;
DR	SMART: SM00109; CL: 1.	RT	"Prediction of the coding sequences of unidentified human genes.XVII.The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."
DR	SMART: SM00324; RhogAP; 1.	RT	DNA Res. 7:143-150(2000);
DR	PROSITE: PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.	RL	EMBL: AB040911; BAB960021; -.
SEQUENCE	632 AA; 71048 MW; TDC06CB725BAAA4 CRC64;	DR	HSSP: Q07960; IRG.
Query Match	99.6%; Score 3330; DB 4; Length 632;	DR	InterPro: IPR002219; DAG_PE-bind.
Best Local Similarity	99.7%; Pred. No. 1.2e-21;	DR	InterPro: IPR00198; RhogAP.
Matches	630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	DR	PROSITE: PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
QY	1 MDTMAMLYNNRNLFEOLVRRVEILSEGNVQFTIQLAKDFEFTRKKWQRTDELGKVKDLRK	DR	Pfam: PF00130; DAG_PE-bind; 1.
Db	1 MDTMAMLYNNRNLFEOLVRRVEILSEGNVQFTIQLAKDFEFTRKKWQRTDELGKVKDLRK	DR	SMART: SM00109; CL: 1.
QY	61 AETERSALDVKLKHARNQDVKEKKRORREADCEKLERQIOLIREMLACDTSSESSIONRE 120	DR	PROSITE: PS50081; DAG_PE_BIND_DOM_1; UNKNOWN_1.
Db	61 AETERSALDVKLKHARNQDVKEKKRORREADCEKLERQIOLIREMLACDTSSESSIONRE 120	FT	NON_TER 1
QY	121 OKSALAFLRGQFSSSNAGNKRISTIDEGSISLSDISKTKTDSLDWMSLWVTFKLRK 180	SEQUENCE	570 AA; 63468 MW; A03F36C0AAC65485 CRC64;
Db	121 OKSALAFLRGQFSSSNAGNKRISTIDEGSISLSDISKTKTDSLDWMSLWVTFKLRK 180	Query Match	90.0%; Score 2919; DB 4; Length 570;
QY	121 OKSALAFLRGQFSSSNAGNKRISTIDEGSISLSDISKTKTDSLDWMSLWVTFKLRK 180	Best Local Similarity	99.8%; Pred. No. 1.9e-19;
Db	121 OKSALAFLRGQFSSSNAGNKRISTIDEGSISLSDISKTKTDSLDWMSLWVTFKLRK 180	Matches	569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	241 YWTRSRRTGTQPNNSDSTLSNROLEPRTEDSVGTPOSNGMRLHDFFSKTVIKPESC 300	QY	63 TERSALDVKLKHARNQDVKEKKRORREADCEKLERQIOLIREMLACDTSSESSIONRE 122
Db	241 YWTRSRRTGTQPNNSDSTLSNROLEPRTEDSVGTPOSNGMRLHDFFSKTVIKPESC 300	Db	1 TERSALDVKLKHARNQDVKEKKRORREADCEKLERQIOLIREMLACDTSSESSIONRE 122
QY	301 VPGCKRKRIGKLSLKCRDCRVSHPECDRCPCLPCITPLTIGTPVKIGEGLADFLVSOPSP 360	QY	123 SALALRNQOPSSSNAGNERLSTIDESGSTLSDFKDKESTLDWMSLWVTFKLRK 182
Db	301 VPGCKRKRIGKLSLKCRDCRVSHPECDRCPCLPCITPLTIGTPVKIGEGLADFLVSOPSP 360	Db	61 SALALRNQOPSSSNAGNERLSTIDESGSTLSDFKDKESTLDWMSLWVTFKLRK 120
QY	421 LKDFLRLNRKLEPLTFRNRAFMEAETIDEDNSIAAMYQAVGELPQANRDTLAFLMHL 480	QY	183 RRSRSRQFVGPPGVKKTRSIGSAVDQGNESTIVAKTTVVPNDGGPIRAVSTIEVW 242
Db	421 LKDFLRLNRKLEPLTFRNRAFMEAETIDEDNSIAAMYQAVGELPQANRDTLAFLMHL 480	Db	121 RRSTSRSRQFVGPPGVKKTRSIGSAVDQGNESTIVAKTTVVPNDGGPIRAVSTIEVW 242
QY	481 ORYAQSPHKKMDYANLAKVFGPTVAHAVPNPDPVTMSODIKRQPKVVERLISPLEWWS 540	QY	303 CGKRKIGKLSLKCRDCRVSHPECDRCPCLPCITPLTIGTPVKIGEGLADFLVSOPSP 362
Db	481 ORYAQSPHKKMDYANLAKVFGPTVAHAVPNPDPVTMSODIKRQPKVVERLISPLEWWS 540	Db	243 TRSRRTGTQPNNSDSTLSNROLEPRTEDSVGTPOSNGMRLHDFFSKTVIKPESC 300
QY	541 QFMVVEQENIDPLHVIENNSAFASTPOTDVKVSLIGPVTPEHQLLKTPSSSSLSQRVRS 600	Db	181 TRSRRTGTQPNNSDSTLSNROLEPRTEDSVGTPOSNGMRLHDFFSKTVIKPESC 240
Db	541 QFMVVEQENIDPLHVIENNSAFASTPOTDVKVSLIGPVTPEHQLLKTPSSSSLSQRVRS 600	QY	363 PSTVWHCVNNEIEQRLTETGLYRISGCDRTVKELKEFLRKVTPVPLLSKVDDTHAIICSL 422
QY	601 TLTKNTPRFGSKSATNGLRQGNFFASPMLK 632	Db	301 PSTVWHCVNNEIEQRLTETGLYRISGCDRTVKELKEFLRKVTPVPLLSKVDDTHAIICSL 360
Db	601 TLTKNTPRFGSKSATNGLRQGNFFASPMLK 632	QY	423 KDFLRLNRKLEPLTFRNRAFMEAETIDEDNSIAAMYQAVGELPQANRDTLAFLMHL 482
RESULT 4	Q9P250 PRELIMINARY; PRM; 570 AA.	Db	361 KDFLRLNRKLEPLTFRNRAFMEAETIDEDNSIAAMYQAVGELPQANRDTLAFLMHL 482
ID	Q9P250 PRELIMINARY; PRM; 570 AA.	QY	483 VAQSPHTKMDYANLAKVFGPTVAHAVPNPDPVTMSODIKRQPKVVERLISPLEWWS 542
AC	01-OCT-2000 (TREMblrel. 15, Created)	Db	421 VAQSPHTKMDYANLAKVFGPTVAHAVPNPDPVTMSODIKRQPKVVERLISPLEWWS 542
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)	QY	543 MMVFEQENIDPLHVIENNSAFASTPOTDVKVSLIGPVTPEHQLLKTPSSSSLSQRVRLST 602
DT	01-JUN-2002 (TREMblrel. 21, Last annotation update)	Db	481 MMVFEQENIDPLHVIENNSAFASTPOTDVKVSLIGPVTPEHQLLKTPSSSSLSQRVRLST 540
DE	KIAA1478 protein (Fragment).	QY	541 TNKTPRFGSKSATNGLRQGNFFASPMLK 570
GN	KIAA1478.	Db	541 TNKTPRFGSKSATNGLRQGNFFASPMLK 570
RESULT 5	Q9BZ74 PRELIMINARY; PRM; 628 AA.	QY	603 TNKTPRFGSKSATNGLRQGNFFASPMLK 602
ID	Q9BZ74 PRELIMINARY; PRM; 628 AA.	Db	541 TNKTPRFGSKSATNGLRQGNFFASPMLK 570

AC	0QBZ74;	ID	Q9WVM1	PRT:	628 AA.
DT	01-JUN-2001 (TREMBrel. 17, created)	AC	Q9WVM1;		
DT	01-JUN-2001 (TREMBrel. 17, last sequence update)	DT	01-NOV-1999 (TREMBrel. 12, Created)		
DE		DT	01-NOV-1999 (TREMBrel. 12, last sequence update)		
GN	FKSG42.	DE	RAC GTPase-activating protein (GTPase activating protein ID-GAP) (MGCRAcGAP variant beta) (MGCRAcGAP variant alpha) (RAC)		
OS	Homo sapiens (Human)	DE	GTPase-activating protein (GTPase activating protein ID-GAP) (MGCRAcGAP variant beta) (MGCRAcGAP variant alpha) (RAC)		
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	DR	GTPase-activating protein 1.		
OX	NCBI_TaxID=9606;	GN	RACgap1.		
RN	[1]	OS	Mus musculus (Mouse)		
RP	SEQUENCE FROM N.A.	OC	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; OX		
RA	Wang Y.-G.; Gong L.;	RA	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
RT	"Characterization of FKSG42, a novel gene located on human chromosome 12.";	RA	NCBI_TaxID=10090;		
RL	submitted (JAN 2001) to the EMBL/GenBank/DBJ databases.	RN	[1]		
DR		RP	SEQUENCE FROM N.A.		
DR	HSSP; Q07960; IRGp;	RA	Medline=99249188; PubMed=10235109;		
DR	InterPro; IPR000198; Rhogap; 1.	RA	"Identification and characterization of a transcript for a novel Rac GTPase-activating protein in terminally differentiating 3T3-L1 adipocytes";		
DR	SMART; SM00324; Rhogap; 1.	RT	RT		
SO	SEQUENCE 628 AA; 70013 MW; A142855A613154CC CRC64;	RT	DNA Cell Biol. 18:265-273(1999).		
Query Match	Best Local Similarity 87.0%; Score 2781; DB 4; Length 628; Matches 550; Conservative 31; Mismatches 47; Indels 4; Gaps 2;	RP	[2]		
QY	1 MDTMMLNVRNLFEQLYRVLVEILSEGNEVQFTOLAKDFEDFRKKWQRTDHELGKYDLM 60	RA	Kawashima T., Kitamura T.;		
Db	1 MDTMMLNVRNLFEQLYRVLVEILSEGNEVQFTOLAKDFEDFRKKWQRTDHELGKYDLM 60	RA	"Mus musculus cDNA for GTPase activating protein McRacGAP, complete cds.";		
QY	61 AETERSALDVKLKHARNQDVIEIKRORADECCEKLERQOLIREMLMCPTSGSTOLSE 120	RA	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.		
Db	61 AETERSALDVKLKHARNQDVIEIKRORADECCEKLERQOLIREMLMCPTSGSTOLSE 120	RT	"Mice with a homozygous gene trap vector insertion in McRacGAP die during pre-implantation development.";		
QY	121 QKSALAFLNQGPSSSNAGNKRSLIDESGSILSDISFDKIDESLDWDSSLSVLUVKFLKK 180	RT	RT		
Db	121 QKSALAFLNQGPSSSNAGNKRSLIDESGSILSDISFDKIDESLDWDSSLSVLUVKFLKK 180	RP	SEQUENCE FROM N.A.		
QY	181 EKRSTSRSQFVGPPGPKVKTRSIGASAVDQGNEISAVKTIVVNPNDGPPIEAUSTETVP 240	RA	Strausberg R.;		
Db	181 EKRCSSESRQFVGPPGPKVKTRSIGASAVDQGNEISAVKTIVVNPNDGPPIEAUSTETVP 240	RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
QY	241 YWRSRRKGTTQPNNSDTSNROLEPRETDVSGTPQSNQGMRLHDFVERTKVPESC 300	RA	Francis A., Baekelandt V., Kozak C.A., Zerial M., Huybrechts D.		
Db	241 YWRSRRKGTTQPNNSDTSNROLEPRETDVSGTPQSNQGMCLNDVSKTVKPEFY 300	RT	"Mice with a homozygous gene trap vector insertion in McRacGAP die during pre-implantation development.";		
QY	301 VPGCKRKFGKLSIKERDCRVTSHSPCCRDRPLPCPTLGTGPVKIGEGGMADFSQTS 360	RT	RT		
Db	301 VPGCKRKFGKLSIKERDCRVTSHSPCCRDRPLPCPTLGTGPVKIGEGGMADFSQTS 360	RP	SEQUENCE FROM N.A.		
QY	361 MIPSTIVWHCVNIEIQQLTETGLYRISGCCDRTVKELKEFLAVKTVPLLSKVDDHHAICS 420	RA	Strasberg R.;		
Db	361 MIPSTIVWHCVNIEIQQLTETGLYRISGCCDRTVKELKEFLAVKTVPLLSKVDDHHAICS 420	RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
QY	421 LKKDFTRNLKEPLTTERLNRAFMEAABITDENNSTIAMYQAVGELPQANRDTLAFMLHL 480	RA	Francis A., Baekelandt V., Kozak C.A., Zerial M., Huybrechts D.		
Db	420 LKKDFTRNLKEPLTTERLNRAFMEAABITDENNSTIAMYQAVGELPQANRDTLAFMLHL 479	RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
QY	481 ORVAQSPHTKMDVANLAQKVEPTIVAHVPRDPVMSQDTRKPKVEREULSPLWEYS 540	RA	Francis A., Baekelandt V., Kozak C.A., Zerial M., Huybrechts D.		
Db	480 ORVAQSPHTKMDVANLAQKVEPTIVAHVPRDPVMSQDTRKPKVEREULSPLWEYS 539	RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
QY	541 QPMVVEOENIDPLVVIENSARSTPQTPDIKVLSSLLGPVTPHQLIKTPSSSSLSORVRS 600	RA	Francis A., Baekelandt V., Kozak C.A., Zerial M., Huybrechts D.		
Db	540 QFSIVEQEKDPLPHVTEENSAFS--TPDKVSLGPTVTPHQLIKTPSSSSLSORVRS 596	RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
QY	601 TLKNTPRFGSKSKSATNLGIQGNFRASPMLK 632	RA	Francis A., Baekelandt V., Kozak C.A., Zerial M., Huybrechts D.		
Db	597 TLKNTPRFGSKSKSATNLGIQGNFRASPMLK 628	RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
Query Match	Best Local Similarity 84.3%; Score 2734; DB 11; Length 628; Matches 534; Conservative 84.4%; Pred. No. 1e-18; Mismatches 47; Indels 6; Gaps 3;	RP	SEQUENCE FROM N.A.		
QY	1 MDTMMLNVRNLFEQLYRVLVEILSEGNEVQFTOLAKDFEDFRKKWQRTDHELGKYDLM 59	RA	Medline=99249188; PubMed=10235109;		
Db	1 MDTMMLNVRNLFEQLYRVLVEILSEGNEVQFTOLAKDFEDFRKKWQRTDHELGKYDLM 60	RA	"Identification and characterization of a transcript for a novel Rac GTPase-activating protein in terminally differentiating 3T3-L1 adipocytes";		
QY	60 60 KAETERSALDVKLKHARNQDVIEIKRORADECCEKLERQOLIREMLMCPTSGSTOLSE 119	RT	DNA Cell Biol. 18:265-273(1999).		
Db	61 KAETERSALDVKLKHARNQDVIEIKRORADECCEKLERQOLIREMLMCPTSGSTOLSE 120	RT	DNA Cell Biol. 18:265-273(1999).		
QY	120 EOKSALAFLNQGPSSSNAGNKRSLIDESGSILSDISFDKIDESLDWDSSLSVLUVKFLKK 179	RT	DNA Cell Biol. 18:265-273(1999).		
Db	121 EOKSALAFLNQGPSSSNAGNKRSLIDESGSILSDISFDKIDESLDWDSSLSVLUVKFLKK 180	RT	DNA Cell Biol. 18:265-273(1999).		

OY	180 REKRSISRSQFVQGPGCQVKKRSIGSVQDGQESIVAKTTVNPNDGGPIEAUSTTIV	239	OY	457 AMYCAVGLPQANDTIAFLMIHQVRQAOSPHTMVDANLAKYEGFTIVAHAVENPDVT	516
Db	181 REKRSISRSQFVQGPGCQVKKRSIGSVQDGQESIVAKTTVNPNDGGPIEAUSTTIV	240	Db	80 AMYCAVGLPQANDTIAFLMIHQVRQAOSPHTMVDANLAKYEGFTIVAHAVENPDVT	139
OY	240 PYWTRSRKGKTLQPNNSDSTLNRSQLEPRTEIDSGVTPQSNGMRLIDFVSKVIRE	299	OY	517 MSQDIIKROPKVVEREULSPLPELYNSQFMAYEQENIDPLAVIENSASIPOTPUKVSLG	576
Db	241 PSWTRSRKGKTLQPNNSDALNSRPLPRTDIDNLGPQNTGMRLLIDFVSKVIRE	300	Db	140 MLQDIKROPKVVEREULSPLPELYNSQFMAYEQENIDPLAVIENSASIPOTPUKVSLG	199
OY	300 CVPCKRKFRKGKLSLKCRDCRLVSHSPECDRCLPCLPCIPPLVGIPVKLGEGMLADFSQS	359	OY	577 PWTPEHOLKTSSSSLSQRVSESTLTKPREGSSKSATNLGRQGFNFSMLK	632
Db	301 CVPCKRKFRKGKLSLKCRDCRLVSHSPECDRCLPCLPCIPPLVGIPVKLGEGMLADFSQS	360	Db	200 PWTPEHOLKTSSSSLSQRVSESTLTKPREGSSKSATNLGRQGFNFSMLK	255
OY	350 PMPSIVWVCNEIEQRLTEGLYRISGCDRVKELKEKELVKTVLISKVDIHC	419			
Db	361 PMPAIVSVCNETEQRGLTEAGLYRISGCDRVKELKEKELVKTVLISKVDIHC	420			
OY	420 SLIKDFLRNLKEPLLTFLRNLRAFMEAETDEEDNSIAMQAVGELPOANRDTLAFLMH	479	RESULT 7	091607 PRELIMINARY; PRT; 625 AA.	
Db	421 SLIKDFLRNLKEPLLTFLRNLRAFMEAETDEEDNSIAMQAVGELPOANRDTLAFLMH	480	ID 091607 AC	091607 DT	
OY	480 LORVAQSEHTKMDYQVANLAKVFGPTIVAHAVPNPDVMTQSODIKRQPKVVEREULSPLTW	539	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	01-MAY-2000 (TREMBLrel. 20, Last annotation update)	
Db	481 LQRVSQSDTAKMDIANLAKVFGPTIVAHAVPNPDVMTQSODIKRQPKVVEREULSPLTW	540	DE CG13345 protein	DE CG13345 protein	
OY	540 SDPMVEQENIDLHLVINSNAESTPOPDIVSLLGVTTPHOLIKTPSSSSLQVR	599	OS Drosophila melanogaster (Fruit fly)	OS Drosophila melanogaster (Fruit fly)	
Db	541 NORMVDOENIDS---ORGNGNSTPRDPDVKLSPVTPFOLVTKPLSSLSQRLY	596	OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pierigyoata; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epipydroidea; Drosophilidae; Drosophila.	OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pierigyoata; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epipydroidea; Drosophilidae; Drosophila.	
OY	600 STIKKNPRFGSSKSATNLGRQGFNSPMLK	632	RN NCBI_TaxID=4227;	RN [1]	
Db	597 N-LSKSTPRFGNKSATNLGRQGFNSPMLK	628	RP SEQUENCE FROM N.A.	RP STRAIN-BERKELEY;	
			RX MEDLINE-20190006; Published=10731132;	RX	
			RA Adams M.D., Gelnicker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	RA	
			RA Adams M.D., Gelnicker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	RA	
			RA Amannatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	RA	
			RA Amannatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	RA	
			RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	RA	
			RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	RA	
			RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Chamep M., Pfleiffer B.D.,	RA	
			RA Wan K.H., Doyle C., Baxter C.G., Heit G., Nelson C.R., Miklos G.L.G.,	RA	
			RA Abrial J.F., Agbavany A., An H.-J., Andrews-Pfaunckoch C., Baldwin D.,	RA	
			RA Bailley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	RA	
			RA Beeeson K.Y., Benos P.V., Belmont B.P., Bhandari D., Bolshakov S.,	RA	
			RA Borckova D., Botchan M.R., Bouck J., Brokshtain P., Brottier P.,	RA	
			RA Burtsis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	RA	
			RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	RA	
			RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	RA	
			RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	RA	
			RA Durbin R.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	RA	
			RA Fosler C., Gabrielian A.E., Garg N.S., Geibelman P., Brottier P.,	RA	
			RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	RA	
			RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,	RA	
			RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	RA	
			RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	RA	
			RA Kimmeil B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	RA	
			RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,	RA	
			RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,	RA	
			RA Merkulov G., Milshina N.V., Mobarry C., Morris J.J., Msahafri A.,	RA	
			RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	RA	
			RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacble J.M.,	RA	
			RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	RA	
			RA Reinhert K., Remington K., Saunders R.D.C., Scheeler F., Shen P.,	RA	
			RA Shue B.C., Siden-Kiamios I., Simpson M., Skupski M.P., Smith T.,	RA	
			RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	RA	
			RA Svartkas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,	RA	
			RA Wang Z.-Y., Wassamman D.A., Weinstein G.M., Weissenbach J.,	RA	
			RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	RA	
			RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	RA	
			RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,	RA	
			RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., Venter J.C.,	RA	
			"The genome sequence of <i>Drosophila melanogaster</i> ";	RT	
			Science 287:2185-2195(2000);	RL	
			EMBL: AE003815; AAF58324.1.;	DR	
			FlyBase: FBgn0033881; RacGAP.	DR	
			InterPro: IPR002219; DAG_PE-bind.	DR	

K08E3.6	protein.	Db	597	ATDCRHAMTALEFEDDVYWFQFLGTS-----VSMASNOIETARHQDNFALCDRSIIG	649
Chernhabditis elegans		RN	577	PVTTPHEHOLLKTPSSSSLSQLQRSTLTKNTPRFGS	611
Eukaryota; Metazoa; Nematoidea; Rhabditida; Rhabditoidea;		RX		:: :	
Rhabditidae; Peloderaidae; Caenorhabditis;		RN	650	PVT-----SPATPLLARSANATARGAHLGS	677
[1]	SEQUENCE FROM N.A.				
RA	MCMURRAY, R.A.,	RA	RESULT 11		
RL	Submitted (NOW-1996) to the EMBL/GenBank/DDBJ databases.	Q95VRL	PRELIMINARY;	PRT;	354 AA.
RN	[2]	ID	Q95VRL		
RP	SEQUENCE FROM N.A.	AC	Q95VRL;		
RX	MEDLINE-99059613; PubMed=9851916;	DT	01-DEC-2001 (TREMBLrel. 19, Created)		
RA	none;	DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
RT	*Genome sequence of the nematode <i>C.elegans</i> : A platform for	DE	Virus-like particle protein VIP2.		
RL	investigating biology';	OS	Venturia canescens.		
Science 282:2012-2018(1998).		OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
DR	EMBL; ZB1568; CAB0453; 1.	OC	Ichneumonoidea; Ichneumonidae; Campopieginae; Venturia.		
DR	InterPro; IPR00198; RhogAP.	OC	NCBI-TaxID=32260;		
DR	PROSITE; PS00479; DAG_PEG_BIND; 1.	OX	[1]		
DR	Pfam; PF00620; RhogAP; 1.	RN			
DR	SMART; SM00109; C1; 1.	SEQUENCE FROM N.A.			
DR	SMART; SM0024; RhogAP; 1.	RA	"Sequence analysis and expression of a virus-like particle protein,		
DR	PROSITE; PS00479; DAG_PEG_BIND_DOM; 2; 1.	RT	VLP2, from the parasitic wasp Venturia canescens";		
SQ	SEQUENCE: PS00479; DAG_PEG_BIND_DOM; 1; UNKNOWN_1.	RL	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.		
Query Match	Best Local Similarity 16.9%; Score 547.5; DB 5; Length 681; Matches 181; Conservative 26.0%; Pred. No. 3.7e-29; Mismatches 284; Indels 121; Gaps 22;	DR	InterPro; IPR00198; RhogAP; -.		
QY	7 NVRNLFEQLVRRVELLSEGEQEVQFQLAKDFEDFRKKW-----QRTDHELGKWDLM 59	PFAM	InterPro; IPR00198; RhogAP; 1.		
Db	14 NSRHFNMILNSQRFQFDIKDIGHFHLDIEIRKLWLKSEESKKRNLADMRPEEA 73	DR	Reinke A.; Asgari S.; Ma G.; Beck M.; Schmidt O.;		
QY	60 KAETERSALDKVLUHARNHOVDVEERRRQAERADCEKLERQIQLTREMLMCMTSSQIQLSE 119	PFAM	Neoptera; Endopterygota; Hymenoptera; Apocrita;		
Db	74 KARKKLMAMEDIDVKDQTQKHRLALMEENKAALKLDDNLYREKQDQKDAKMKNGIFNS-LIK 131	DR	Ichneumonoidea; Ichneumonidae; Campopieginae; Venturia.		
QY	120 EOKSALAFLNRGQOPSSSNAGNKRLSTIDESGSTLSDISKDFTDLSLMD---SSLVKTFK 176	PFAM	NCBI-TaxID=32260;		
Db	132 EDROQPKFKE--PVLRTYSKRQV-QTHPHMELDTQDDEDDSEFVDYETGDSFEEVH 186	DR	[1]		
QY	177 LKK-REKRSTS-----RQFDQPPG----PVKK 200	DR	VLP2, from the parasitic wasp Venturia canescens";		
Db	187 LRNGREVRSSAAGNAVGGRKSASAHAITAANSKRSRSRVMATIDEPNEGGTPPKR 246	EMBL	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.		
QY	201 TRSISGAVDQGNESIVAKITVT----VENDGP-----IEAVSTIETWP 240	EMBL	InterPro; IPR00198; RhogAP; -.		
Db	247 CRDGSGTPHQEMTTTTTTTINNSRAQNQDPFRVSLHQQLTRSLSCGSPSCDQTP 306	PFAM	Reinke A.; Asgari S.; Ma G.; Beck M.; Schmidt O.;		
QY	241 YWTRSRRKIGTGQFWNSDSTINSQLEPRPTEDSVGTPOUNGSNHRLHDENVKTPKESC 300	DR	Neoptera; Endopterygota; Hymenoptera; Apocrita;		
Db	307 GOTNNINIGGMSSALTIKSTDLDIRTLKRGTPAWNGTTR-DIAMRPTTIEAGKAMRK 365	PFAM	Ichneumonoidea; Ichneumonidae; Campopieginae; Venturia.		
QY	301 VPGCKRKIFGKGLSLKCRCRVSHPECRDRCPICP-----TLLGTPVKGEM 350	DR	NCBI-TaxID=32260;		
Db	366 DKCATALKLA-TSMKCRDCHQVHHSNCNKLHLPCCIPRPTMMTPKSALRGAKPGAGEFR 424	EMBL	VLP2, from the parasitic wasp Venturia canescens";		
QY	351 LADEVSQTPSPMISIVVHVQNELEFQRLGILTETGLYRURISGQDRTVWEEKFLRVKVNPILL 410	PFAM	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.		
Db	425 LQDPTCTSKPMPIRAVHCVVALEARGLTQEGIYFVPGQVRTVWILDE-LRSKTVPNV 483	DR	InterPro; IPR00198; RhogAP; 1.		
QY	411 KVDDHATCSSLKDFLRLNKEPLFLRMRFAEMAEL-TDEDNISTIAAMYOAVGELPOA 468	EMBL	Reinke A.; Asgari S.; Ma G.; Beck M.; Schmidt O.;		
Db	484 -LHDVEVITDLKRFRLDKDPLPRTSRQELTVANLSTDPOINGRALRNVICELPOA 542	PFAM	Neoptera; Endopterygota; Hymenoptera; Apocrita;		
QY	469 NRDTPLAFMLTHIQRY-AOSPHTKMDVANLAKVGFPTIVAHAVPNPDVMSQ-----D 520	DR	Ichneumonoidea; Ichneumonidae; Campopieginae; Venturia.		
Db	543 NRDTPLAYLFTHWRKVIAQSSRNKNCAMARMVAPAVMGH-----PVQSQSOAIGRD 596	EMBL	InterPro; IPR00198; RhogAP; 1.		
521 IKRQPKVVERLLSILEWSQFMWAEQEMIDPVIWLIENSNAFSTPQTPD---IKVSLIG 576	PFAM	Reinke A.; Asgari S.; Ma G.; Beck M.; Schmidt O.;			
QY	SEQUENCE FROM N.A.	DR	Reinke A.; Asgari S.; Ma G.; Beck M.; Schmidt O.;		
RC	STRAIN=11S, AND ISS.	EMBL	Neoptera; Endopterygota; Hymenoptera; Apocrita;		

QY 259 STLNRSQLEPRTEDSVGTPQSNGMLRLDFVSKWIKPBCSCVPGKTRKGKLSL-LKC 316
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 31 ATIKENEYHPKYE-----KVNFKVHTFRGFWCECA-NFMWGLIAGQVKC 76
 QY 317 RDCRVSHPECRCPCLPCIPYL-----IGTPVKIGEMMADFYSOTSPPHPSIV 366
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 77 ADCGGLNVRHQCSKWWPNDOCKPDLLKKVVKVYSCDLTTLV-----AHFTRR----PMV 125
 QY 367 VHCNVNEIQRGLITETGLYRISGCDRTKELKEKLUR-VKTVPPLSKVDDIHACSLKD 424
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 126 DMCTIREIESRGLNSEGLYLYRVSFGSFLDIEPVKMAFDGKADISVNMYEDINITGALKL 185
 QY 425 FLRNKKEPLTFLRNRAFMEAETTDNSIAAMYQAVGELPOANRDTIAFLMIHLQRYA 484
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 186 YFROLPIPLITYDAPKFESAKTVDPPDQELEIHEALKLPPACETLYMLMAHLKRV- 244
 QY 485 QSPHTR--MDVANLAKVFGPTIVAHAVNPDPTMSQIKRQPRKVERIL 532
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 245 -TLHEKENLMSAENLGIYGPTLM--RAPELDAMALNDIYORLIVVEMLI 292
 QY 427 -----NPRTSPPLASNAE-----NKPPVQAQQSPPVLLPLPPIOTTTQS 467
 Db 291 SKTVKIKESCVCQGKRIKFGKLSLKCRDRVSHPEGDRCP---LCICP---TLIGT- 342
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 427 -----NPRTSPPLASNAE-----NKPPVQAQQSPPVLLPLPPIOTTTQS 467
 QY 343 -----PVKIGCMLDFVSGS-----PMPSIVVHCV 370
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 468 REVAPPSSINNSNRAASPFRTSVPSSPSPITSKSLLEGARLDAILREHSNPTNMQCT 527
 QY 371 NETEQRGLITETGLYRISGCDRTKELKEKLURVTKVPL-----SKVDDIHACSLKD 424
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 528 SOVENFGNLNOGIYRVSSARVNMLRSQF--ENNPLQLHTPEDYENDHAWADLUKI 584
 QY 425 FLRNKKEPLTFLRNRAFMEAETTDNSIAAMYQAVGELPOANRDTIAFLMIHLQRYA 484
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 585 FFRELREPLIPDNHORDFDAGNEVEDSRRDVAHRAINDLDPDANYSTIRHLTHLAKIK 644
 QY 485 Q-SPHTRMDVANLAKVFGPTIVAHAVNPDPTMSQIK 522
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 645 ENSDVNMNSTNLLAITWGPTLTIQAT-IPETTSFSRTE 682
 QY 094466 PRELIMINARY; PRY; 695 AA.
 ID 094466;
 AC 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 CC -1-SIMILARITY: TO OTHER EUKARYOTIC RHO GDP-DISSOCIATION INHIBITORS.
 DR EMBL; AL035065; CAA22624.1; -.
 DR HSSP; 007960; 1RGP.
 DR InterPro; IPR01060; Cdc15_Fes_C1P4.
 DR Pfam; PF00611; FCH; 1.
 DR Pfam; PF00620; RhogAP; 1.
 DR SMART; SM0055; FCH; 1.
 DR SMART; SM00324; RhogAP; 1.
 KW Hypothetical protein; GTPase activation.
 FT DOMAIN 426 POLY-SER.
 SQ SEQUENCE 695 AA; 76678 MW; IDEAF1328D576263 CRC64;
 QY Query Match 9.5%; Score 308; DB 3; Length 695:
 Best Local Similarity 23.7%; Pred. No. 9.2e-13;
 Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23;
 Oy 35 KDFEDRKWKWORDHE---LGKVKKLMAATE-RSALDKLKHARNQDVETK--RR 86
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 145 KSTKEYAKKQENAYLEAVNQMDKSKRSRKGAAETQYRNALDKN----NTGDSQKRYGEEK 199
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Qy 87 ORAEDCEKLERQIQLIREMLMCDSGSOLSEE-QKSALAF-----LNRGOPS 134
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 200 PKSNAQLTQKLEBARRKALENAESDMHSKTEAQNWQKOLCIHRRPYNIKOFFSQREIES 259
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Qy 135 SSNAGNKRULSTIBDEGSILSDI-----SFDKTDPSLDMDSSLVTTKFEKLK 178
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 260 SLIANYLRYTKCCESTNLINGLTIRPQKPTPTNGLQHALDNTNANTDQVOYVLA-SIR 318
 Qy 179 KREKRST--SRQFVGPGPVKTRSTGSAVDOGNESTIVAKTWTVPNGGPE---A 232
 ::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Db 319 HEQNKNPDTASKRKKIQPSSYGCNGSSACKTNPPNPTI-KVTAIRAS--PQONTNP A 373



On protein - protein search, using sw model
Run on : July 3, 2003, 09:44:06 ; Search time 73 Seconds
(without alignments)
1153.622 Million cell updates/sec

Title: US-09-881-736-2
Perfect score: 3243
Sequence: 1 MDTMLNVRNLPEQLVRRVE.....SKSATNLGROGNFFASPMLK 632
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcdata/geneseq/geneseq/geneseq-emb1/AJ1980.DAT: *
2: /SIDS2/gcdata/geneseq/geneseq/geneseq-emb1/AJ1981.DAT: *
3: /SIDS2/gcdata/geneseq/geneseq/geneseq-emb1/AJ1982.DAT: *
4: /SIDS2/gcdata/geneseq/geneseq/geneseq-emb1/AJ1983.DAT: *
5: /SIDS2/gcdata/geneseq/geneseq/geneseq-emb1/AJ1984.DAT: *
6: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1985.DAT: *
7: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1986.DAT: *
8: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1987.DAT: *
9: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1988.DAT: *
10: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1989.DAT: *
11: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1990.DAT: *
12: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1991.DAT: *
13: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1992.DAT: *
14: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1993.DAT: *
15: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1994.DAT: *
16: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1995.DAT: *
17: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1996.DAT: *
18: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1997.DAT: *
19: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1998.DAT: *
20: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ1999.DAT: *
21: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ2000.DAT: *
22: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ2001.DAT: *
23: /SIDS2/gcdata/geneseq/geneseq-emb1/AJ2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3233	99.7	632 22	AAM40047 Human polypeptide
2	3230	99.6	632 23	AAU69419 Lung small cell ca
3	2919	90.0	570 23	AAU69422 Lung small cell ca
4	1182	36.4	255 22	AAB94223 Human protein sequ
5	971.5	30.0	256 22	AAU32069 Novel human secret
6	864	26.6	625 22	ABB68056 Drosophila melanog
7	625	19.3	4318 22	ABG14787 Novel human diagno
8	618	19.1	665 22	ABG10230 Novel human diagno
9	618	19.1	1086 22	ABG19804 Novel human diagno
10	19.1	1139 22	ABC13958 Novel human diagno	

RESULT 1
ID AAM40047
AA standard; Protein: 632 AA.
AC AAM40047;
XX DT 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 3192.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; KW leukaemia.
XX OS Homo sapiens.
XX PN WO200153312-A1.

XX PD 26-JUL-2001.
XX PR 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0652191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.

ALIGNMENTS

Line	Score	Length	DB ID	Description	Line	Score	Length	DB ID	Description
11	581	17.9	22	AAM41833 Human Polypeptide	11	581	17.9	22	AAM41833 Human Polypeptide
12	477	14.7	22	ABR59247 Drosophila melanog	12	477	14.7	22	ABR59247 Drosophila melanog
13	322	9.9	22	AAM39659 Human polypeptide	13	322	9.9	22	AAM39659 Human polypeptide
14	322	9.9	22	AAM41445 GTPase activating	14	322	9.9	22	AAM41445 GTPase activating
15	301	9.3	1261	19 AAY9268 Human G-protein ac	15	301	9.3	1261	19 AAY9268 Human G-protein ac
16	301	9.3	22 AAB97911 Human G-protein ac	16	301	9.3	22 AAB97911 Human G-protein ac		
17	294.5	9.1	193	22 AAB67553 Amino acid sequenc	17	294.5	9.1	193	22 AAB67553 Amino acid sequenc
18	290.5	9.0	736	20 AAW97809 Human GTPase regul	18	290.5	9.0	736	20 AAW97809 Human GTPase regul
19	290.5	9.0	759	22 ABG16793 Novel human diagno	19	290.5	9.0	759	22 ABG16793 Novel human diagno
20	290.5	9.0	800	22 ABG16794 Novel human diagno	20	290.5	9.0	800	22 ABG16794 Novel human diagno
21	286.5	8.8	856	21 AAB41660 Novel signal trans	21	286.5	8.8	856	21 AAB41660 Novel signal trans
22	286.5	8.8	1173	22 AAB97911 Novel human diagno	22	286.5	8.8	1173	22 AAB97911 Novel human diagno
23	278.5	8.6	967	22 ABG12188 Sequence encoded b	23	278.5	8.6	967	22 ABG12188 Sequence encoded b
24	270.5	8.3	589	7 AAB60303 Human myosin Ix	24	270.5	8.3	589	7 AAB60303 Human myosin Ix
25	270.5	8.3	2548	20 AAY05781 Novel human diagno	25	270.5	8.3	2548	20 AAY05781 Novel human diagno
26	267	8.2	1212	22 ABB20843 Drosophila melanog	26	267	8.2	1212	22 ABB20843 Drosophila melanog
27	265	8.2	494	22 ABB59717 RhoGAP domain from	27	265	8.2	494	22 ABB59717 RhoGAP domain from
28	263	8.1	555	22 AAU17313 Novel human diagno	28	263	8.1	555	22 AAU17313 Novel human diagno
29	263	8.1	985	22 ABB12190 Novel human diagno	29	263	8.1	985	22 ABB12190 Novel human diagno
30	260.5	8.0	802	20 AAY21700 Novel human diagno	30	260.5	8.0	802	20 AAY21700 Novel human diagno
31	258	8.0	903	21 AAB42926 Novel human diagno	31	258	8.0	903	21 AAB42926 Novel human diagno
32	256.5	7.9	659	23 ABB51281 Novel human diagno	32	256.5	7.9	659	23 ABB51281 Novel human diagno
33	251	7.7	170	22 AAG65821 Novel human diagno	33	251	7.7	170	22 AAG65821 Novel human diagno
34	251	7.7	22 AAG67552 Novel human diagno	34	251	7.7	22 AAG67552 Novel human diagno		
35	249	7.7	2110	22 ABB58077 Novel human diagno	35	249	7.7	2110	22 ABB58077 Novel human diagno
36	248.5	7.7	526	22 AAB93466 Novel protein seq	36	248.5	7.7	526	22 AAB93466 Novel protein seq
37	247.5	7.6	291	22 AAU17333 Novel signal trans	37	247.5	7.6	291	22 AAU17333 Novel signal trans
38	246.5	7.6	296	22 AAM25391 Human protein seq	38	246.5	7.6	296	22 AAM25391 Human protein seq
39	246.5	7.6	718	22 ABB66573 Drosophila melanog	39	246.5	7.6	718	22 ABB66573 Drosophila melanog
40	246	7.6	816	22 AAG66505 GTP enzyme Rho fam	40	246	7.6	816	22 AAG66505 GTP enzyme Rho fam
41	241.5	7.4	1844	22 ABB58723 Drosophila melanog	41	241.5	7.4	1844	22 ABB58723 Drosophila melanog
42	238	7.3	974	23 ABBG18722 Prostate cancer-as	42	238	7.3	974	23 ABBG18722 Prostate cancer-as
43	232	7.2	322	22 ABBG66699 Novel human diagno	43	232	7.2	322	22 ABBG66699 Novel human diagno
44	231	7.1	818	23 ABB90757 Human Tumour Endöt	44	231	7.1	818	23 ABB90757 Human Tumour Endöt
45	229.5	7.1	668	22 ABBG27731 Novel human diagno	45	229.5	7.1	668	22 ABBG27731 Novel human diagno

XX (HYSEQ) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB; AAI59203.

XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 4; SEQ ID NO 3192; 10078pp; English.

CC The invention relates to human nucleic acids (AAI57798-AA161369) and
 CC the encoded polypeptides (AAW38642-AAW42213) with nootropic,
 CC immunosuppressant and cytosolic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activation/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 632 AA:
 SQ Query Match 99.7%; Score 3233; DB 22; Length 632;
 Best Local Similarity 99.7%; Pred. No. 6.8e-263; Mismatches 0; Gaps 0;
 Matches 630; Conservative 0; Indels 2; Deletions 0;

QY 1 MDTMMLNVARNLFEQLYRREVILSESGNEVQFIQLAKDFERPKKKWQRTDHELGKYKDLMK 60
 Db 1 MDTUMMLNVNHLFEQLYRREVILSESGNEVQFIQLAKDFERPKKKWQRTDHELGKYKDLMK 60
 QY 61 AETERSALDVQLKHARNQDVKEKRRQABADCENTERIQLREMLMDTSGSIOLSE 120
 Db 61 AETERSALDVQLKHARNQDVKEKRRQABADCENTERIQLREMLMDTSGSIOLSE 120
 QY 122 QKSALAFLURGKGOPESSNACKRKSTIDESGSISLSDIFKTDESUDWOSLVKFKLRR 180
 Db 122 QKSALAFLURGKGOPESSNACKRKSTIDESGSISLSDIFKTDESUDWOSLVKFKLRR 180
 QY 181 EKRSTSROFVGPGPVKTRSTGSAVQGENESTIVAKTVTVNDGGIEAVNSTIETP 240
 Db 181 EKRSTSROFVGPGPVKTRSTGSAVQGENESTIVAKTVTVNDGGIEAVNSTIETP 240
 QY 241 YWRSRKGTLQWNSDTSLSQLEPRETTSVGPGSNGGRLHDTSKTIKPEC 300
 Db 241 YWRSRKGTLQWNSDTSLSQLEPRETTSVGPGSNGGRLHDTSKTIKPEC 300
 QY 301 VPCGKRKIKGKLKCRDRVVSHECRCPICPTIGTPVKGEMLADFRVSQHSP 360
 Db 301 VPCGKRKIKGKLKCRDRVVSHECRCPICPTIGTPVKGEMLADFRVSQHSP 360
 QY 361 MIPSIWHRVNETEQRGLTETGLYRISGDRTYELKEFLRVTKVPLSKVDDIHACS 420
 Db 361 MIPSIWHRVNETEQRGLTETGLYRISGDRTYELKEFLRVTKVPLSKVDDIHACS 420
 QY 421 LKADFLRNKKEPLTFRMRAFMMAETDEDDNSIAMQAVGLPQARDTIAFLMH 480
 Db 421 LKADFLRNKKEPLTFRMRAFMMAETDEDDNSIAMQAVGLPQARDTIAFLMH 480
 QY 481 ORVAQSPHFRMDVANLARVFGPTIVAHAVPNPDPVTMSODIKRQPKVVERLISPLEWWS 540
 Db 481 ORVAQSPHFRMDVANLARVFGPTIVAHAVPNPDPVTMSODIKRQPKVVERLISPLEWWS 540

RESULT 2
 ID AAU9419 standard; Protein: 632 AA.
 XX AAU9419
 AC AAU9419;
 DT 30-JAN-2002 (first entry)
 XX DE Lung small cell carcinoma antigen #13.
 XX PR Human; cytostatic; antitumour; lung small cell cancer antigen;
 KW tumour; lung cancer.
 XX OS Homo sapiens.
 XX PR WO200171768-A2.
 PN XX PD 18-OCT-2001.
 XX PF 11-APR-2001; 2001WO-US11859.
 PR 11-APR-2000; 2000US-195780P.
 PR 21-JUN-2000; 2000US-213361P.
 PR 01-SEP-2000; 2000US-221763P.
 PR 05-SEP-2000; 2000US-236629P.
 PR 14-SEP-2000; 2000US-232565P.
 PR 19-DEC-2000; 2000US-257037P.
 PR 08-JAN-2001; 2001US-260796P.
 PR XX PA (CORI-) CORIXA CORP.
 PI Loddes MJ, Wang T, Mohamath R, Indirias CY;
 PR XX DR WPI: 2002-010896/01.
 PR N-PSDB; AAS61859.
 XX PT Lung tumour polynucleotide and polypeptides useful in therapy and
 XX diagnosis of cancer especially lung cancer
 PS Claim 2; Page 270-271; 295pp; English.
 XX The invention relates to novel isolated lung small cell cancer antigen
 CC polynucleotides (I) and polypeptides (II) used in a method of detecting
 CC cancer in a patient. The method is optionally performed by
 CC utilising oligonucleotides (III), where the biological sample
 CC from the patient is contacted with (III), detecting the amount of
 CC polynucleotide hybridised to (III) in the sample and comparing the
 CC amount of polynucleotide to a predetermined cut-off value and thereby
 CC determining cancer in a patient. (I), (II) or antigen-presenting cells
 CC expressing (II) is useful for stimulating and/or expanding T cells
 CC specific for a tumour protein. The method comprises contacting T cells
 CC with one of the components under conditions to permit the stimulation
 CC and/or expansion of the cells. A composition comprising (I) is useful for
 CC stimulating an immune response in a patient and for inhibiting the
 CC development of a cancer especially lung cancer in a patient. An
 CC isolated T cell population is useful for removing tumour cells from the
 CC biological sample and for inhibiting the development of cancer in a
 CC patient. AAU9407-AAU9431 represent novel human lung small cell
 CC cancer antigen amino acid sequences of the invention.
 XX Sequence 632 AA;
 SQ Query Match 99.6%; Score 3230; DB 23; Length 632;

	Best Local Similarity 99.7%; Pred. No. 1.2e-262; Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	PR	14-SEP-2000; 2000US-232565P.
QY	1 MDTMALNRLNEFOLVVRVEILSEGNEQFIQLAKDFEDFRKKWQRTDHELGKYKDLMR	PR	19-DEC-2000; 2000US-25037P.
DP	61 AEVERSALDVKLKHARNQDVDETRRQAEADCEKLERQIQLTREMILCDTSQIOLSEEE	PR	08-JAN-2001; 2001US-260796P.
QY	61 AEVERSALDVKLKHARNQDVDETRRQAEADCEKLERQIQLTREMILCDTSQIOLSEEE	XX	(CORI-) CORIKA CORP.
DB	121 OKRALAFLENLRGQSSNSNAGNKRISTIDSGSILDISDKTDISLDMSSLVTFKIKR	XX	Lodes MJ, Wang T, Mohamath R, Indrias CY;
QY	121 OKSLAFLAFLNRGQSPSSNAGNKRUSTIDSGSILSHISFDKTDESLDWSSLVKFKLRR	XX	DR WPT: 2002-010895/01.
DP	181 EKRSTSOVDFOPPGPKTRSIGSADQGENESTIVAKTTVTPNDGGPIEAVSTIEVW	XX	DR N-PSDB; RAS61862.
QY	181 EKRSTSOVDFOPPGPKTRSIGSADQGENESTIVAKTTVTPNDGGPIEAVSTIEVW	PT	Lung tumour polynucleotide and polypeptides useful in therapy and diagnosis of cancer especially lung cancer
QY	241 YWTRSRRTGTGLQPWNSTLNRSQLEPRTEDSVGTQDQSNGGRHLDFVSKTVIKPESC	PT	Claim 2; Page 282-284; 295pp; English.
DP	241 YWTRSRRTGTGLQPWNSTLNRSQLEPRTEDSVGTQDQSNGGRHLDFVSKTVIKPESC	XX	XX
QY	301 VPGCKRKIFGKLSLKCRGCRVSHPECRDRCPUCIPICPLIGTPVKIGGMLAFLVFSQSP	CC	The invention relates to novel isolated lung small cell cancer antigen polynucleotides (I) and polypeptides (II) used in a method of detecting cancer in a patient. The method is optionally performed by utilising oligonucleotides (III), where the biological sample from the patient is contacted with (III) detecting the amount of polynucleotide hybridised to (III) in the sample and comparing the amount of polynucleotide to a predetermined cut-off value and thereby determining cancer in a patient. (I), (II) or antigen-presenting cells expressing (II) is useful for stimulating and/or expanding T cells specific for a tumour protein. The method comprises contacting T cells with one of the components under conditions to permit the stimulation and/or expansion of the cells. A composition comprising (I) is useful for stimulating an immune response in a patient and for inhibiting the development of a cancer especially lung cancer in a patient. An isolated T cell population is useful for removing tumour cells from the biological sample and for inhibiting the development of cancer in a patient. AAU69407-AAU65431 represents novel human lung small cell cancer antigen amino acid sequences of the invention.
DB	301 MISIIVWCVNELEQRGLTETGLYRISCDRTKELKELKFRLKVTKVPLISKVDIATCS	CC	CC
QY	360 361 MISIIVWCVNELEQRGLTETGLYRISCDRTKELKELKFRLKVTKVPLISKVDIATCS	CC	CC
DP	421 LLKDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL	CC	CC
QY	421 LLKDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL	CC	CC
DB	481 QRYAQSPHTKMDYANLAKVFGPTIVAHAVPNPDVPTMQLDIKQPKVVERLISPLEWS	CC	CC
QY	481 QRYAQSPHTKMDYANLAKVFGPTIVAHAVPNPDVPTMQLDIKQPKVVERLISPLEWS	CC	CC
DP	541 QFMVNEQNIIDPHVHVIENSAFEPOTDVKISLUGPTTPEHOLKPPSSSLSQRRS	CC	CC
QY	541 QFMVNEQNIIDPHVHVIENSAFEPOTDVKISLUGPTTPEHOLKPPSSSLSQRRS	CC	CC
DB	601 TLTKNTPRGSKSKSATNLGROQNEFAASPLMK	XX	XX
QY	601 TLTKNTPRGSKSKSATNLGROQNEFAASPLMK	XX	XX
DP	601 TLTKNTPRGSKSKSATNLGROQNEFAASPLMK 632	XX	XX
RESULT 3		Sequence	570 AA:
AAU69422	Query Match 90.0%; Score 2919; DB 23; Length 570; Best Local Similarity 99.8%; Pred. No. 1.5e-236; Matches 569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	QY	183 RRSTSROFVGPPGPKTRSIGSADQGENESTIVAKTTVTPNDGGPIEAVSTIEVW
ID	AAU69422 standard; Protein: 570 AA.	DB	121 RRSTSROFVGPPGPKTRSIGSADQGENESTIVAKTTVTPNDGGPIEAVSTIEVW
XX		QY	121 RRSTSROFVGPPGPKTRSIGSADQGENESTIVAKTTVTPNDGGPIEAVSTIEVW
AC	AAU69422;	DB	243 TRSRRTGTGLQPWNSTLNRSQLEPRTEDSVGTQDQSNGGRHLDFVSKTVIKPESC
XX		QY	243 TRSRRTGTGLQPWNSTLNRSQLEPRTEDSVGTQDQSNGGRHLDFVSKTVIKPESC
DT	30-JAN-2002 (first entry)	DB	181 SALAFLNRCQSPSSNAGNRLSTIDESGSIISLSDSFKDIDESLDWSSLVTFKIKR
DE	Lung small cell carcinoma antigen #16.	QY	181 SALAFLNRCQSPSSNAGNRLSTIDESGSIISLSDSFKDIDESLDWSSLVTFKIKR
XX	Human; cytosatic; antitumour; lung small cell cancer antigen; KW tumour; lung cancer.	DB	61 SALAFLNRCQSPSSNAGNRLSTIDESGSIISLSDSFKDIDESLDWSSLVTFKIKR
XX	Homo sapiens.	QY	61 SALAFLNRCQSPSSNAGNRLSTIDESGSIISLSDSFKDIDESLDWSSLVTFKIKR
OS		DB	120 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
XX		QY	120 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
PW	W020017168-A2.	DB	361 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
XX		QY	361 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
PD	18-OCT-2001.	DB	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
XX		QY	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
PF	11-APR-2001; 2001WO-US11859.	DB	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
XX		QY	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
PR	11-APR-2000; 2000US-196780P.	DB	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
PR	21-JUN-2000; 2000US-213361P.	QY	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
PR	01-SEP-2000; 2000US-229763P.	DB	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL
PR	05-SEP-2000; 2000US-230629P.	QY	423 KDFRNKLEPKLTERINRAFMEAETDEDSIAAMQAVGELPQANRDTLAFLMTHL

QY 543 MAVEQENIDPLHVIENSNAFSTPQTDIKVSLIGVTPPEHQQLKTPSSSLSORVRSTL 602
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 481 MMEGENIDPLHVIENSNAFSTPQTDIKVSLIGVTPPEHQQLKTPSSSLSORVRSTL 540
 QY 603 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 541 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 570
 QY 397 KEKEFRVKTVPLLSKVDDTAICSLKDTRNLKEPLTFLRNIAFMEAETDEDNSIA 456
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 20 KLKFLLRKVTKVPLLSKVDDTAICSLKDTRNLKEPLTFLRNIAFMEAETDEDNSIA 79
 AC AAB94223 standard; Protein; 255 AA.
 XX
 AC AAB94223;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:14587.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 XX
 OS
 XX EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/4.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs.
 XX
 PS Claim 8; SEQ ID 14587; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAB03166 to AAB13628 and
 CC AAB11842 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAB13632
 CC represent oligonucleotides, all of which are used in the exemplification
 XX
 SQ Sequence 255 AA;

Query Match 36 4%; Score 1182; DB 22; Length 255;
 Best Local Similarity 99 2%; Pred. No. 6e-91; 2; Indels 0; Gaps 0;
 Matches 234; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 457 AMYQVGEPLQANRDTLAELMHIQORVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPV 516
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 80 AMYQVGEPLQANRDTLAELMHIQORVAQSPHTKMDVANLAKVFGPTIVAHAVPNPDPV 139
 QY 517 MSQDITKRPKVVERILSPLPEWSPOMMVEQENIDPLHVIENSNAFSTPQTDIKVSLG 576
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 140 MLQDKRQPKVERISLSPLEYWSPFMVEQENIDPLHVIENSNAFSTPQTDIKVSLG 199
 QY 577 PVTTEPHOLIKTPSSSLSORVRSTLTKNTPRGSKSKSATNLGRQGNFFASPMLK 632
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 200 PVTTEPHOLIKTPSSSLSORVRSTLTKNTPRGSKSKSATNLGRQGNFFASPMLK 255
 QY 543 MAVEQENIDPLHVIENSNAFSTPQTDIKVSLIGVTPPEHQQLKTPSSSLSORVRSTL 602
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 481 MMEGENIDPLHVIENSNAFSTPQTDIKVSLIGVTPPEHQQLKTPSSSLSORVRSTL 540
 QY 603 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 632
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 541 TKNTPRFGSKSKSATNLGRQGNFFASPMLK 570
 QY 397 KEKEFRVKTVPLLSKVDDTAICSLKDTRNLKEPLTFLRNIAFMEAETDEDNSIA 456
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 20 KLKFLLRKVTKVPLLSKVDDTAICSLKDTRNLKEPLTFLRNIAFMEAETDEDNSIA 79
 AC AAB94223 standard; Protein; 255 AA.
 XX
 AC AAB94223;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:14587.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 XX
 OS
 XX EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI: 2001-318749/4.
 XX
 PT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #2560.
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haemopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia;
 KW
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PR 25-OCT-2001.
 XX
 PT 16-APR-2001; 2001WO-US086656.
 XX
 PR 18-APR-2000; 2000US-0552929.
 PR 26-JAN-2001; 2001US-0770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy -
 XX
 PS Claim 20; Page 551; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation, to regulate haemopoiesis, and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and

CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
SQ Sequence 256 AA;

Query Match 30.0%; Score 971.5; DB 22; Length 256;
Best Local Similarity 84.0%; Pred. No. 3.1e-73; Matches 205; Conservative 3; Mismatches 27; Indels 9; Gaps 4;
Db 1 QTPSPMPSIVHCVNEIQRGTEGTYKLRISGDRYKKFLRVKTVPLISKVDDIH 60
QY 357 ATCSLLKDFRLNKEPLTFLRNLRAFMMAETDEDNTIAAMQAVELPQANRDTAFL 416
Db 417 ATCSLLKDFRLNKEPLTFLRNLRAFMMAETDEDNTIAAMQAVELPQANRDTAFL 476
QY 61 ATCSLLKDFRLNKEPLTFLRNLRAFMMAETDEDNTIAAMQAVELPQANRDTAFL 120
Db 477 MHILORYAQSPTRKMDYANLAKVFGPTVYAHAVNPUPPTMSQDIR-OPKVRLISLP 535
Db 121 MTHLQRVQAQSPHTKMDYANLAKVFGPTVYAHAVNPUPPTMSQDIR-OPKVRLISLP 180
QY 536 L-EWWSPEMVEE-- NIDPLAVIENNSAFSTPQDPDIKV--SLIGPVITPERHOLLK 587
Db 181 LWEYWESEFMDWEQGTFDPHYIENNSAATPQDPIKAVPGGICWHTAGEELQK 240
QY 588 TRSS 591
Db 241 GPPS 244

RESULT 6
ABB68056
ID ABB68056 standard; Protein: 625 AA.
XX
AC ABB68056;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 30960.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
DR N-PSDB; ABL12159.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -
XX
PS Disclosure: SEQ ID NO 30960; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions - useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL3051), expressed DNA
CC

CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABL5737-ABL72072).
CC The sequence data for this patent did not form part of the printed
CC Specification, but was obtained in electronic format directly from WIPO
CC at ftp://wipo.int/pub/published_pct_sequences.

XX Sequence 625 AA;
SQ Sequence 625 AA;
Query Match 26.6%; Score 864; DB 22; Length 625;
Best Local Similarity 35.1%; Pred. No. 1.5e-63; Matches 228; Conservative 110; Mismatches 225; Indels 86; Gaps 21;
Db 1 MAL5ALASFDLRRCMQVLTDPPEERFLRFMEQYHEKAGYAETARIQNEDKSL 60
QY 63 TERDALDVKHARNQDVIEKRRQRAEADCKLERIOQLREMLMDTSSEISIOLBEEK 122
Db 61 TMKGDLGKLFHARRILDEMIKARRQAEHERADMESKIMAYADLRRHEN---LNNETR 116
Db 123 SALAFLNRQGQPSS---NA--GNKLSTIDESGLSISDFDKI-DESLDWDSIYKT 174
Db 117 DKLAFLHT-LPSSRKRSKLNAVREDKSYGDINSTGSLSDLSITHSDEDFD---VRT 170
QY 175 FKIKKREKRSTSRQFV-----DGPPGPVKTR--SIGAVD---- 209
Db 171 SK-SWREHRPSLKPKNQIPSGVKRSLRSTGLNGMSCTTPGKSRASSVGIGEVHTD 229
QY 210 -OGNESIVAKTTVTENDG-GPEIAVSTIERTVYWPYRSTSRRKTGTQWPWNSDSTLNSRQL 266
Db 230 VSGGAERFCAATTKVITPODGQVIRAMSTIESLPIVAGNERGGDSLSTPRRSVLKEATA 289
QY 267 EPTETTSV-----GTP-OSNGGHLHDVSKTWIKPESCPGCKRIFGKLSKCR 317
Db 350 DCYRCHIDCRTLTVSCVPO-TGTPTKTMGYVTDFASTPAMPALVHCVNELETR 349
Db 290 PPLTPVNAMAPHHVAEStCPLQHRLMRNHTFSQTKFLRGDNVCQQCKRIRGAVGLRCR 349
QY 318 DCRVSVHPECRDRCLPCLCPILTGTP-VKIGEMLADFVSQTSPMPSIVHCVNETEQR 376
Db 350 DCYRCHIDCRTLTVSCVPO-TGTPTKTMGYVTDFASTPAMPALVHCVNELETR 408
Db 377 GLUTETGLYKRISSGDRYKKFLRVLKELKEPLTFLRNLKEPLT 436
Db 409 GLTEVGLYRISSEREVKALKQFLRKATPHIGNT-DIVLCCVCKDFLRSITEFLIPT 467
QY 437 RLNRFAEMAETDEDNTIAAMQAVELPQANRDTAFLMLHCVNELETR 496
Db 468 SQWKFDFANAVQNDPDKTAQDMFLVKSQQLPQANRDTAFLLHFQRIAQCPVWLMPIDN 527
QY 497 ARVFGPTVYAHAVNPUPPTMSQDIRKQPKVVERLSSLPLWWSQMFANEGENIDPLHVI 556
Db 528 SLIGFTPTIVGYSIPPDQDHAITYTEVFKOMVKAELLPVSPWEOYIV---IDPTR-- 580
QY 557 ENSNAFSTPOT---PDKVSLLGPVTP-----EHQLKTPSSSS 593
Db 581 -----TATVIKRVPSNKNLDSLYATPFKGKTIKKRFVGTTPASA 622

RESULT 7
ABG14787
ID ABG14787 standard; Protein: 4318 AA.
XX
AC ABG14787;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #14778.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.

XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 XX DR Dmanac RT, Liu C, Tang YT;
 PI XX WPI; 2001-639362/73.
 DR N-PSDB; AAS78974.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
 PT Claim 20: SEQ ID NO 45146; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 4318 AA:
 CC
 PS Query Match 19.3%; Score 625; DB 22; Length 4318;
 PS Best Local Similarity 38.8%; Pred. No. 4.3e-42; Mismatches 167; Conservative 46; Indels 132; Gaps 16; Matches 3915 QY 216 VAKTIVTVVPNDGGPPE-----AVSTIEVTPWTRSR-----RKTGIVQ----- 253
 Db 3915 IAKSILSQKRNKAGGITLPPFKLYVKATVVKTTWYQHRAVDQWNRGSEIMPHINYL 3974
 QY 254 -----PRNSTLNS-----ROLE-----PRTEIDS-----VG 276
 Db 3975 IFDKPKKNKRGKDSLNFNKCWNWLAICRKKLKNPLPTYIKINSRWKDLHVRPKTIK 4034
 QY 277 TPOSNGGMRUH-----DVSXKTVKPECVCPCKRKIGKLISKCRDRVSVHECRRL 330
 Db 4035 TLEENLGNITRQVIGMGRDNMSKT----PKAMATKAKIDKBLDLIKSFFC----- 4080
 QY 331 CPLPCIPITLIGTPKIG-----EGMLADAFVQSPTSPMISIVVHCVNNEEQ----- 375
 Db 4081 -----TKETTITRVNRPTEWEKIFAYSSD-----KGLSIRVNELKQIVKKTNNP 4128
 QY 376 -----RGLTERGLY-----RISGCDR-----TVEBLKEKFLRKYTPVLP----- 409
 Db 4129 IKKWFNDMMRHFSKEDIYAKKHKMKCSSLPAIREMOQTITMRYHLTPVRIALIKSGNN 4188
 QY 410 SKVDDIHACCSLLKDFLRNLKEPLTRNRAFNAEAETDEDNSIAAMQAVGELPQAN 469
 DB 4189 SKVDDIHACCSLLKDFLRNLKEPLTRNRAFNAEAETDEDNSIAAMQAVGELPQAN 4248

RESULT 8
 ID ABG10230 standard; Protein; 665 AA.
 XX DE Novel human diagnostic protein #10221.
 XX PN WO200175067-A2.
 XX PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 XX PI Dmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 DR N-PSDB; AAS74417.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
 PT
 XX PS Claim 20: SEQ ID NO 40589; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 665 AA:
 CC
 PS Query Match 19.1%; Score 618; DB 22; Length 665;

QY 470 RDTLAFMLIHQLQRAQSPTKMDVANLAKVFGPTIVAHAVPNPDFTMSQDIKRQPKVVE 529
 Db 4249 RDTLAFMLIHQLQRAQSPTKMDVANLAKVFGPTIVAHAVPNPDFTMSQDIKRQPKVVE 4308
 QY 530 RLSSPLEYW 539
 Db 4309 RLSSPLEYW 4318

Best Local Similarity 46.9%; Pred. No. 8.1e-43; Matches 153; Conservative 35; Mismatches 68; Indels 70; Gaps 11;

QY 255 WNSDSLNSRQLRERTEIDSVGTPQSNGMRLHDVFVSKVIKEBCVCGKRTFGKJSL 314
Db 369 WIKD--LNRPKTITKTLERNLGNITQDGK--KDFMSKT--PKAMATTKIKWDLQI 422
CC 315 KCRDCRVVSHPECRDRCPCLPCIPNLIGPVKIG-----EGMLADFVSOTSPMPSTVH 368
QY 423 KSFFC-----TAKEETIRVNQOPTEWEKIFATYSSD----KGLISR 459
Db 369 CYNIEQ-----RGJETGLX---RISGDR--TYKELKEK-FLYV 403
QY 460 TYNELKOIYKKTNPIKKWTDNMRHFSKEDTYAAKKHMCKSSLPAIREMOQTTHY 519
Db 404 KTVPLL-----SKVDIHAICSLIKFLRNKLKEPLTFLRNRAFMEAETDDN 453
Db 520 HLTVPRIATIKKSGNNNSKDIHAICSLIKFLRNKFPLTFLRNRAFMEAETDDN 579
QY 454 SIAMMQAVGELPOANRDTLAFMLIHQRVAQSFHTKDVANLAKVFGPTIVAHAVPND 513
Db 580 SISAMQAVGELPOANRDTLFLMLHQVRQAQSFYKVMVANLAEVFGSTIVAHAVPNE 639
QY 514 PVTM5ODIKRQPKYVERLLSLPLXYW 539
Db 640 PVTMLOQDIKQPKYVERLPSLPLXYW 665

RESULT 9

ID ABG19904 standard; Protein: 1086 AA.

AC ABG19904;

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #19895.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PP 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

N-PSDB; AAS84091.

PT New isolated polynucleotide and encoded polypeptides, useful in PT diagnostics, forensics, gene mapping, identification of mutations PT responsible for genetic disorders or other traits and to assess PT biodiversity -

XX Claim 20; SEQ ID NO 50263; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and CC poly peptide (II) sequences (I) is useful as hybridisation probes, CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC .and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags CC for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity.

The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences.

Sequence 1086 AA:

Query Match 19.1%; Score 618; DB 22; Length 1086;

Best Local Similarity 46.9%; Pred. No. 1.8e-42; Matches 153; Conservative 35; Mismatches 68; Indels 70; Gaps 11;

QY 255 WNSDSLNSRQLRERTEIDSVGTPQSNGMRLHDVFVSKVIKEBCVCGKRTFGKJSL 314
Db 790 WIKD--LNRPKTITKTLERNLGNITQDGK--KDFMSKT--PKAMATTKIKWDLQI 843
QY 315 KCRDCRVVSHPECRDRCPCLPCIPNLIGPVKIG-----EGMLADFVSOTSPMPSTVH 368
Db 844 KSFFC-----TAKEETIRVNQOPTEWEKIFATYSSD----KGLISR 880
QY 369 CYNIEQ-----RGJETGLX---RISGDR--TYKELKEK-FLYV 403
Db 881 TYNELKOIYKKTNPIKKWTDNMRHFSKEDTYAAKKHMCKSSLPAIREMOQTTHY 940
QY 404 KTVPLL-----SKVDIHAICSLIKFLRNKLKEPLTFLRNRAFMEAETDDN 453
Db 941 HLTVPRIATIKKSGNNNSKDIHAICSLIKFLRNKFPLTFLRNRAFMEAETDDN 1000
QY 454 SIAMMQAVGELPOANRDTLAFMLIHQRVAQSFHTKDVANLAKVFGPTIVAHAVPND 513
Db 1001 SISAMQAVGELPOANRDTLFLMLHQVRQAQSFYKVMVANLAEVFGSTIVAHAVPNE 1060
QY 514 PVTM5ODIKRQPKYVERLLSLPLXYW 539
Db 1061 PVTMLOQDIKQPKYVERLPSLPLXYW 1086

RESULT 10

ID ABG13958
AC ABG13958 standard; Protein: 1139 AA.

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #113949.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PP 11-OCT-2001.

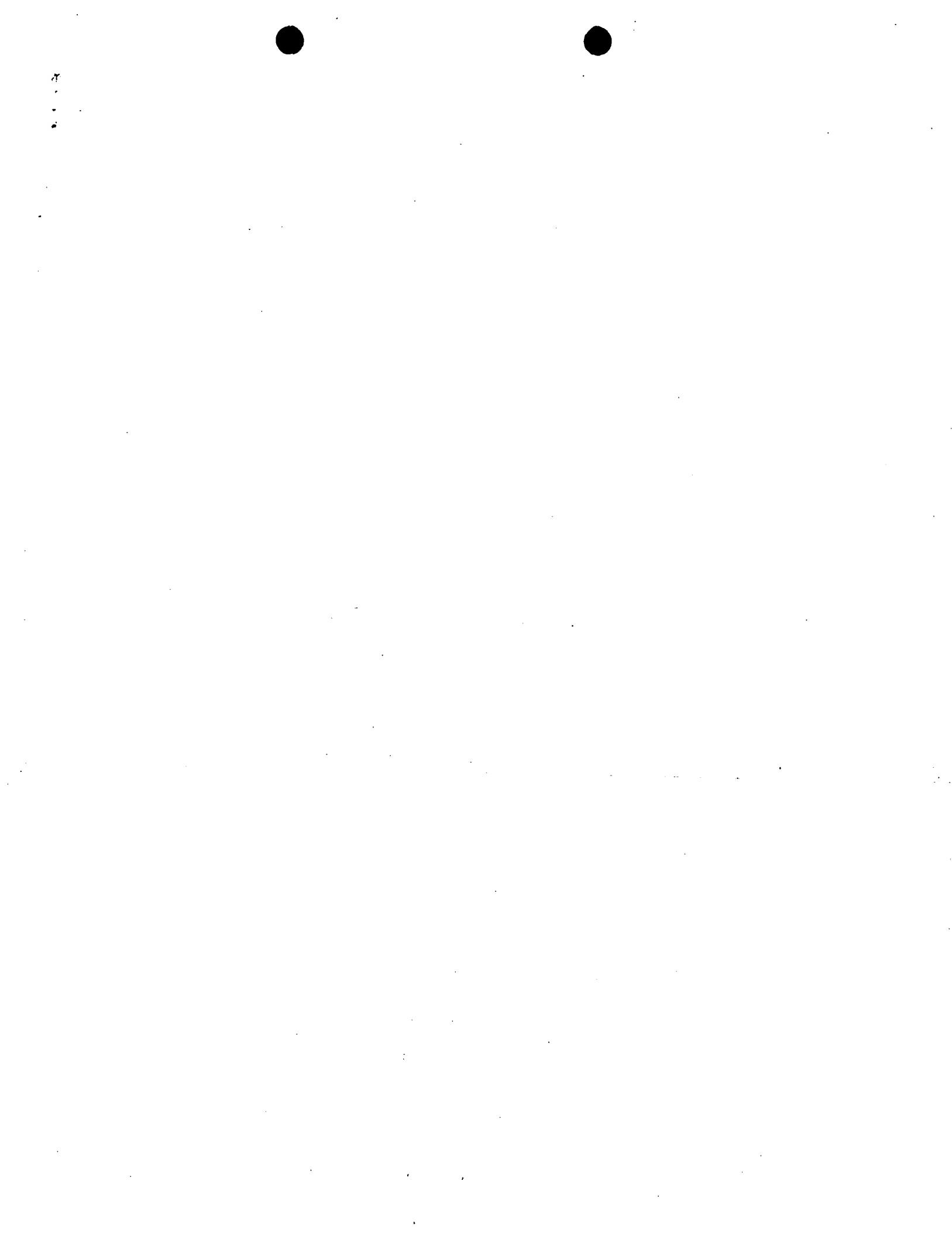
XX 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX	WPI; 2001-639352-773.	DR	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX	N-PSDB; AAS78145.		
PT	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification or mutations responsible for genetic disorders or other traits and to assess biodiversity	KW	peripheral nervous system; neuropathy; central nervous system; CNS;
PT		KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
PT		KW	amyotrophic lateral sclerosis; Shy Drager Syndrome; chemotactic;
PT		KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
PT		KW	leukaemia.
XX	Claim 20; SEQ ID No 44317; 103pp; English.	OS	Homo sapiens.
PS		XX	
XX	The invention relates to isolated polynucleotide (I) and polypeptide chain reaction (PCR) primers, oligomers, and for chromosome polymers, and gene mapping, and in recombinant production of (II). The poly nucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (III) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity.	XX	
CC	The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequence. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.	XX	
CC	Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pat/published_pct_sequences .	XX	
SQ	Sequence 1139 AA;	XX	
Query Match	Best Local Similarity 19.1%; Score 618; DB 22; Length 1139; Matches 158; Conservative 45.0%; Pred. No. 1.9e-42; Indels 84; Gaps 14;	XX	
Qy	242 WTRSRKRGTLQPNNSDST-LNSR----QLEPRTEIDTSVGTPOSSNGMRLH----DF 289	XX	
Db	820 WLAICRK---KLDFPLTAVKINNSRWNKDLNIRKTT---IKTLEENGLTIQDGMGKF 874	XX	
Qy	290 VSKVIKPECPSCVPGKKRFLGKSLCRDCRVCWHPPECRDCPCLPCIPRLIGTPKRIG- 347	XX	
Db	875 MSKT--PKAMATKTKIDKWDLQLQKSFFC-----TAKEETIRVNRO 913	XX	
Qy	348 ---EGMLADFVFSOTSPMPISIVVHCVNIEQ-----RGLFTGTY- 384	XX	
Db	914 PTWEKIFATYSSD---KGLISRYNLKQIYRKRTNNPKWNTNDKMRHFSKEDYA 968	XX	
Qy	385 --RISGDR--TVKELKEK-FLRVKTVLL-----SKVDIHAICSLIKDFLRN 428	XX	
Db	969 AKKHMKKCSLPAIREMGKTTMWHLYTFVRIAITKSKGNNSKVDIHAICSLIKDFLRN 1028	XX	
Qy	429 LKEPLLTFRINKAEMAEAITDEDSNISAMYQAVGELPOANRDTFLMIHLQRQAQSPH 488	XX	
Db	1029 FKEPLLTFRINKAEMAEAITDEDSNISAMYQAVGELPOANRDTFLMIHLQRQAQSPH 1088	XX	
Query Match	Best Local Similarity 17.9%; Score 581; DB 22; Length 119; Matches 115; Conservative 99.1%; Pred. No. 5e-41; Indels 0; Gaps 0;	XX	
Qy	1 MDTMMLNVRNFQEQYQVRVVELSEGNEQFOFLAKOPDFRKWQRTDHELGKYKDLMK 60	XX	
Db	4 MDTMMLSVRNLFQEQYQVRVVELSEGNEQFOFLAKOPDFRKWQRTDHELGKYKDLMK 63	XX	
Qy	61 AETERSALDVKIKHARNQDVDEIKRQRAEADCEKFLRQJLREMIMCDISGSIQ 116	XX	
Db	64 AETERSALDVKIKHARNQDVDEIKRQRAEADCEKLERQJLREMIMCDISGSIQ 119	XX	
RESULT 11		XX	
AM41833	489 TKMDVANIAKVGKPTIVAHAVPPDPVYMSQDITRQPKVVERLISLPLEW 539	XX	
AM41833	1089 TKNVANIAEVFGSTIVAHAVPNPEPVIMLQDIKQPKVVERLISLPLEW 1139	XX	
RESULT 12		XX	
AM41833	AC	AC	
AC	AC	AC	
AC	AC	AC	
DT	22-OCT-2001 (first entry)	XX	
DE	Human polypeptide SEQ ID NO 6764.	XX	
DE	ABB55247	XX	



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

ALIGNMENT

hypothetical prote
 hypothetical prote
 hypothetical prote
 hypothetical prote
 KIA1204 protein [
 RHGDP/LIM domain
 GTPase activat
 hypothetical prote
 RatBPI rat
 KIA0672_protein [
 hypothetical prote
 rata-binding prote
 hypothetical prote
 hypothetical prote
 hypothesis-activator p
 hypothetical prote

QY 481 QRYAQSPKPKMDVANLAKVFGPTIVAHVNPDPVTMSQDIDIKQPKVVERLISPLEWS 540
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T23454
 R;McMurray, A.
 submitted to the EMBL Data Library, November 1996
 A;Reference number: Z19743
 A;Accession: T23454
 A;Status: preliminary; translated from GB/EMBL/IDB
 A;Molecule type: DNA
 A;Residues: 1-681 <WIL>
 A;Cross-references: EMBL:Z81568; PIDN:CAB04593.1; GSPDB:GN00021; CESP:KO8E3.6
 C;Genetics:
 A;Gene: CESP:KO8E3.6
 A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3

RESULT 2

T23454 hypothetical protein KO8E3.6 - Caenorhabditis elegans

QY 469 NRDIAFLIMTHLQRV-AQSPHTKMDVANLAKVFGPTIVAHVNPDPVTMSQ-----D 520
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)
 C;Genetics:
 A;Gene: FlyBase:rn
 A;Cross-references: FlyBase:FBgn0003263

RESULT 3

B48122 Grpase-activating protein Rac homolog, splice form clone pcl.7d - fruit fly (Drosophi

c;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 7 NVRNLFEQDVYRVEILSSECNEVQFIOLAKDFEDPRKKW-----QRTDHELGKYKDLM 59
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 14 NSRHIFNMILNSQRQPQFDKIDGFMHLDEIERLRLKWLWDSEESKKRNLADMRAEEALA 73
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 60 KAFTERSALDVKLKHARQYDVEKRRORAEADCEKLEFQIQLIREMLMDTGSQIOLSE 119
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 74 KARRKLAMPDIDYKTQTKILRALMEENKALKDLUNVYEREREKQLDKAMNGIENS--LTK 131
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 120 EOKSALAFINRGQPSSSNAGNKRULSTIDESGLSILSDISRDKTDESLDWMD--SSLVKTF 176
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 132 EDROQKELRE--PLVRTYSKRIQ--QHRPHMEDTDQDEDSEVDDETGDSPEEWH 186
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 177 LKK--REKRETS-----RQFVDPGGP---PVK 200
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 187 LNRNGREVRRKSAAGNAVGSKRRSSAHAATTAANSKRSRSRVMTATIDEPBEPNGGTPPKR 246
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 201 TRSIGSANDQGNESIVATTV-----VPNDGGP-----IEAVSTIEWP 240
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 247 CRDGGSTPQEMTTTTTHTHNSRQNODPRVSLRSLSGSPSCDQTP 306
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 301 VPGCKRIKFGKLSLKCRDGRVSHPECRCLPCIP-----TLIGTPVKGEGM 350
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 366 DKCATALKLA-TSMKCRDQHVYHRSCCNKHLHLPCLPKRPTMMTPKSALRGAKPGAGER 424
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 351 LADFVQSQTSPMSIVWCVNEEBQRLGETTGLYKRISGCRDITKELKKFLRKTVPLS 410
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similarity 40.7%; Pred. No. 6.5e-23;

Matches 110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;

A;Map position: 3
 A;Introns: 36/1; 73/2; 237/3; 361/3; 612/3
QY 425 LOFQTSKAPMMAVHCVWALRGLTQEGIYRPGVQYVNLDE-LASKTVPHG 483
 C;Species: Drosophila melanogaster
 C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
 C;Accession: B48122
 R;Agnel, M.; Roder, L.; Vola, C.; Griffin-Shea, R.
 Mol. Cell. Biol. 12, 5111-5122, 1992
 A;Title: A Drosophila rotund transcript expressed during spermatogenesis and imaginal
 A;Reference number: A48122; MUID:93024458; PMID:1406685
 A;Accession: B48122
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-383 <AGN>
 A;Note: sequence extracted from NCBI backbone (NCBIN:115662, NCBIPI:115663)

C;Genetics:

A;Gene: FlyBase:rn

A;Cross-references: FlyBase:FBgn0003263

Query Match 14.7%; Score 477; DB 2; Length 383;

Best Local Similar

A;Gene: FlyBase:nr	Qy	252 LQPWNSDTSILNSROLEPRTEITDSVGTPOSQGGMRLHDIFVSKTKVTPKGESCPGKVIKFGK 311	
&;Cross-references: FlyBase:FBgn0003263	Db	: :	
Query Match	Qy	1569 FEPVKGKAQKKRKRQEROVQ-----EING---HVFAASYQVNIPQSOCQLSYTWMD 1617	
Best Local Similarity 14.7%; Score 477; DB 2; Length 384;	Db	312 LSURKDCRVSHPECRDRCPLCPT----LJGTPVYRIGEMLADFSQTSMPM1 365	
Pred. No. 6.5e-23; Mismatches 110; Conservative 44; Indels 14; Gaps 6;	Matches	110; Conservative 44; Mismatches 102; Indels 14; Gaps 6;	
Qy	279 QSNGG-MRLHDIFVSKT-VIKPESCPGKRIKGKLSKMRCDRCRVSHPECRDPCLCI 336		
Db	: :		
78 QSHSLLREBHNFKKISYYVNVGNCWCRIRFAMASLRGRACPIRCHIGCCRQCVN1 137	Db	1618 KALLCSVCKMTCHHKCVH1QSYCISYTRGRKSELGAEPGHF--GCVDSITSKDKAVSPV1 1675	
Qy	337 P-TLIGTPVKGSCMMLADEVSQTSMPM1 336		
Db	: :		
138 PQPQIGTK---RGCLSDYAPRVAFMVPAVLIVHCVTEIEARGLQOEGLYRSSTBRKCKR 193	Db	366 WIVCVNELEQRGLIETGLRISGCDRTVKEELKEPLRKVYKPYLSSKVDD--IAHATCSLK 423	
Qy	396 LKIKFLRKYTKVPLISKVQDIIHATCSSLKQFLRNKEPLIPLRNKAFAEMAEITEDNS1 455		
Db	: :		
194 LERKLRLRKSTPHGN-KHTHTLCCVCKDPLQYHPLIYHRIDFEETARHEDLAVE 252	Db	1676 LEKLLERHEVNGHLXTEGLYRKSGANRTRBLQ---ALQDPATVKLEDPIHAITGVKL 1732	
Qy	456 AAMQVAGELPQANRDTLAFMLMHQIQRVAQSPHKTMDVANLAKVFGPTIVAHAVPNPDV 515		
Db	: :		
253 MAYTVLAVLHLHQAHRTDLAYLMHWKLAESPATVNTVNLAVTAPTLG-----DLD 306	Db	1733 QWLRLPEPLIMTAQDPLQYDFRAMEPLKEPKOBLAQIAYAVIDHLPEAHNTSLERLHLVKY 1792	
Qy	516 TMSQDIKKRQVVERLLSIPLEWNSQMMV 545		
Db	: :		
307 LTLENWVWTQRLVLUVLLMPAGFWNSQFLEV 336	Db	1792 484 A-OSPHIKDVKANVLAKFVFTIVAHAVPNPDVMSQDITRKOPKVERLLSIPLEWNSQF 542	
RESULT 5	Qy	1793 ALLEDVNRMSPGALIAIFACLU-RCPDNISDPLTSMKDVKITCVCVEMLIKEQMKYK-- 1849	
554307	Db	: :	
myosin heavy chain - rat	Db	543 MMVQEOENIDPLHVLEN-----SNA-----FSTQTPDVKSLGPVTPEHQI 585	
C;Species: Rattus norvegicus (Norway rat)	Db	: :	
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 02-Mar-2001	Db	1850 --VKMEEINHLEAESIAFRRLSLRQNAPWPLKLGFSSPIEGVTKSPPTPVQDLEE, 1907	
R;Accession: S54307	Db	586 LKTPSSSS-----LSQRVHS-----TLLKNTPRGSKSATNQGRQGFFFAPI 629	
R;Reinhard, J.; Scheel, A.A.; Diekmann, D.; Hall, A.; Ruppert, C.; Baebler, M.	Db	1908 GALPEEAAGGDEDREKEILMERIOSIKEEREDITYRLPELDPRGSDEENLSETASTES 1967	
A;Title: A novel type of myosin implicated in signalling by rho family Gtpases.	Db	Qy	630 MLK 632
R;Accession number: S54307; MUID:95188874; PMID:7882373	Db	: :	
A;Status: preliminary; nucleic acid sequence not shown	Db	Qy	1968 LLE 1970
A;Molecule type: mRNA	Db		
A;Residues: 1-1980 <REI>	Db		
A;Cross-references: EMBL:X77609; NID:9639998; PIDN:CA54700.1; PID:9539999	Db		
C;Superfamily: rat myosin heavy chain; myosin motor domain homology; protein kinase c zi	Db		
C;Keywords: nucleotide binding; P-loop	Db		
F;149-942/Region: myosin motor domain homology <MMOT>	Db		
F;239-442/Region: nucleotide-binding motif A (P-loop)	Db		
F;1593-1641/Domain: protein kinase C zinc-binding repeat homology <K2N>	Db		
Query Match	RESULT 6		
10.7%; Score 348; DB 2; Length 1980;	554307		
Best Local Similarity 22.1%; Pred. No. 1e-13;	A59256		
Matches 160; Conservative 105; Mismatches 262; Indels 196; Gaps 25;	myosin-1xb [similarity] - human		
Qy	A59256		
82 EKRRQRREADCCKLKERQQLIR-EMLMDTSGSSIQSLSEBEQSKRALAFLNRQGP---SSS 136	C;Species: Homo sapiens (man)		
Db	C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 02-Mar-2001		
1272 QIQRQ-Y-HPTDRLATAWEIWRCKKLASMLSQLSDLSEBKPRAGAAALTPTEERRIFS 1329	C;Accession: A59256; 161700		
Qy	R;Wirth, J.A.; Jensen, K.A.; Post, P.L.; Bement, W.M.; Mooseker, M.S.		
137 NAGKRUSTIDESGSIISLSDIFDK-----TDESIDLWDSSVYTFK 176	J. Cell Sci. 109, 653-661, 1996		
Db	A;Title: Human myosin-1xb, an unconventional myosin with a chimerin-like rho/rac GTP		
1330 TSDVSKLSPVKTSTEVDGDSLASKKPGHHKSSEDPSAGDAGLPPGSGQDSKSAKRLFL 1389	A;Residues: 1-2022 <WIR>		
Qy	A;Reference number: A59256; MUID:97065843; PMID:8907710		
177 LKKREKRSS-----TSQRFVQDP-----PGPVKTRSI 204	A;Cross-references: GB:U42391; NID:9147782; PIDN: AAC50402.1; PID:91147783		
Db	R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.		
1390 HKAKDKKKPSLEGVEETEGSGGQAQEAQAPRKTLDPVSSQHRTGEPLKGKKRNRKV 1449	Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994		
Qy	A;Title: Identification and overlapping expression of multiple unconventional myosin		
235 TIETPV-----YWRSRKRTG 251	A;Reference number: A55758; MUID:94294418; PMID:8022818		
Db	A;Status: preliminary; translated from GB/EMBL/DBJ		
1510 TMISVPNGKIKHVGKKDLMENYQIVSNLAERGERKDNTLVNIVQFSQSLDEFTRSNKT-D 1568	A;Molecule type: mRNA		
Qy	A;Residues: 234-322 <REIS>		
222 LSGRGENQ-----FIQIAKDFEDFRKKWQ--RTDHLIGKVKDLMMAERTSALD 70	A;Cross-references: GB:L129149; NID:9457257; PIDN:AAA20912.1; PID:9531142		
Db	A;Genetics:		
1404 LSQGQVSKSTKRLFLAKTTKD---KRYSLRGAELNNAVGHVWLEATTMKKGLE 1458	A;Gene: GDB:MO9B; OMIM: 602129		
Query Match	Best Local Similarity 23.0%; Pred. No. 8.3e-13;		
1450 GQTIVSERKWRSEVKRITNANELKDELFLINKVNDLSOKTPFESSLEATERRSNIK 1509	A;Map position: 19p13.1		
1450 GQTIVSERKWRSEVKRITNANELKDELFLINKVNDLSOKTPFESSLEATERRSNIK 1509	C;Superfamily: rat myosin heavy chain; myosin motor domain homology; protein kinase		
1390 HKAKDKKKPSLEGVEETEGSGGQAQEAQAPRKTLDPVSSQHRTGEPLKGKKRNRKV 1449	C;Keywords: nucleotide binding; P-loop		
1510 TMISVPNGKIKHVGKKDLMENYQIVSNLAERGERKDNTLVNIVQFSQSLDEFTRSNKT-D 1568	F;149-941/Region: myosin motor domain homology #status atypical <MMO>		
1510 TMISVPNGKIKHVGKKDLMENYQIVSNLAERGERKDNTLVNIVQFSQSLDEFTRSNKT-D 1568	F;239-246/Region: nucleotide-binding motif A (P-loop)		

QY 71 ---KLKHARNQDVKEIKRRQRAEADCEKLERIQOLIREMLMCDTSQGILQSLBHQSKA 126
 : : || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1459 PSGQOHRRHAGE----KRTKEPGGKKRKNRVI-----GKTVSEKWRSSWF 1502
 QY 127 FLNRGOPSSNAGNKRRLSTIDE-SGSIGSDISFDKTDESSLWVKTFLKLRREKR 185
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1503 ---RQITMAN---ELVYDELFLINKINDLRSQKTP-----IESLFIATEKFRS 1545
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 186 TSROFVDRPFPGPVKTRSGSIGAVDOGNESIVAKUTVTVPNDGGPIEAVESTIET-VPMYR 244
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1546 NIKTMYSVNGK1---HVGYKOLMENYQIVVSNLATERGOKDINLVNLVLFQSLDEFTR 1601
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 245 SRRKGTQPLQPWNDSSTLNRSQLEPRTEDSVGTPOQSNQGMRLHDVFVSKTIVKBRSCVFCG 304
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1602 GYTR-----NDPEPVKOSKOSKAQKKRKGKERAVOEHNG---HVEASYCOSIPOSCEDCL 1650
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 305 KRIKFGKLISLKCRCRVVSHPECRDCPCLPCIPL-TGTP-VKIGE-GMLADFSVOTSP 360
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1651 SYVFLMDKALLCSVCKWCKHKKVHKIOSHCSVTYGRKGEGPGRPHGFCVQCVDSLTSKA 1710
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 361 MIPSIIVVHGKVNETOBORGJLTETGLYRISGCDRTYKELKEKLR-VKTVPLISKVDIHACISLKDFLRNKEPLL 418
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1711 SVPVILEKLEHVEMHGLTIEGLIRKSGANRTRRELQ---ALQTDPAVKLENFPIAI 1767
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 419 CSLUKDFURNLKPLTFLRNRAEITDEDNAAMYQAVELPQANRDTLAFMLI 478
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1768 TGVLKQWLRLPELPELMTPTAQYDFPLRAVELPEKEQOLAYAILEHLPEBHNLSRLIF 1827
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 479 HLORVA-QSPHTEMDVANLAKVFGPTIVAHAVENPDPTMSQDIKRQVWVERLPLE 537
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1828 HLVALVEDVNRMSPGALAITAPCIL-RCPDNSDPLTSMKDVLKTTTCVEMILIKEOMR 1886
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 538 YWSOFMMVYEQENIDPLHTEI-----SMAFSTQTPD-----IKVSLLGWT 579
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1887 KYK---VKEMEESOLEAAESTIAFRLSLRQNANKSPKRETAGGAGRLLTTSRVSP-- 1940
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 QY 580 TPHEOLLKUPPSSSLISQRVST-----LTKNPFR 608
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 1941 SPSTRNLALGGSWSAALRTRGTPARPGRARALARRRPF 1980
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 RESULT 7
 S29128
 N-chimera - rat
 N;Alternate names: Gtpase-activating protein
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 25-Feb-1994 #sequence_revision 19-Oct-1995 #text_change 08-Oct-1999
 C;Accession: S29128; S25152
 R;Hall, C.; Monfries, C.; Smith, P.; Lim, H.H.; Kozma, R.; Ahmed, S.; Vanniasingham, J. Mol. Biol. 211:11-16, 1990
 A;Title: Novel human brain cDNA encoding a 34,000 Mr protein n-chimera, related to A;Reference number: S08242; MUID:90133942; PMID:2299665
 A;Molecule type: mRNA
 A;Accession: S08242
 A;Cross-references: EMBL:X51408; NID:935012; PIDN:CAA35759.1; PID:935013
 C;Supfamily: protein kinase C zinc-binding repeat homology <KZ>
 F;46-95/Domain: protein kinase C zinc-binding repeat homology <KZ>
 Query Match
 Best local similarity 9.9%; Score 322; DB 2; Length 299;
 Matches 90; Conservative 48; Mismatches 117; Indels 26; Gaps 8;
 N-chimera - human
 N;Alternate names: Gtpase-activating protein
 C;Species: Homo sapiens (man)
 C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Nov-1999
 C;Accession: S08242
 R;Hall, C.; Monfries, C.; Smith, P.; Lim, H.H.; Kozma, R.; Ahmed, S.; Vanniasingham, J. Mol. Biol. 211:11-16, 1990
 A;Title: Novel human brain cDNA encoding a 34,000 Mr protein n-chimera, related to A;Reference number: S08242; MUID:90133942; PMID:2299665
 A;Molecule type: mRNA
 A;Cross-references: EMBL:X51408; NID:935012; PIDN:CAA35759.1; PID:935013
 C;Supfamily: protein kinase C zinc-binding repeat homology <KZ>
 F;46-95/Domain: protein kinase C zinc-binding repeat homology <KZ>
 Query Match
 Best local similarity 9.9%; Score 322; DB 2; Length 299;
 Matches 90; Conservative 48; Mismatches 117; Indels 26; Gaps 8;
 N-chimera - common canary
 N;Alternate names: protein kinase C homolog [misidentification]; song control circuit
 C;Species: Serinus canaria (common canary)
 C;Date: 26-Aug-1999 #sequence_revision 26-Aug-1999 #text_change 04-Feb-2000
 C;Accession: A43953
 R;George, J.M.; Clayton, D.F.
 Brain Res. Mol. Brain Res. 12, 323-329, 1992
 RESULT 9
 A43953
 N-chimera - common canary
 N;Alternate names: protein kinase C homolog [misidentification]; song control circuit
 C;Species: Serinus canaria (common canary)
 C;Date: 26-Aug-1999 #sequence_revision 26-Aug-1999 #text_change 04-Feb-2000
 C;Accession: A43953
 R;George, J.M.; Clayton, D.F.
 Brain Res. Mol. Brain Res. 12, 323-329, 1992

A;Title: Differential regulation in the avian song control circuit of an mRNA predicting A;Reference number: R43953; MVID:9225621; PMID:1374499	Db	319 HEDNKNPFDASKRKLQOPPSGAGTNPVNPTI-KVTAIPS---PLQNTNP
A;Status: preliminary	Oy	233 VSTIETPYWYTSRRRTGTQLQWNSDSTLN--SROLEPRTEIDSVGTPQSNGMRLHDFV
A;Molecule type: mRNA	Db	374 PSTDPEPNPSVASPAFPNSSTSPTAPASAPLASTNPSTANDTNSSSSS
A;Residues: 1-299 <GEO>	Oy	291 SKIVIKFESCVCGKRRKGFLSLKRCRDRCVVSHPEDCRDPC--IPCIP---TLIGT
A;Experimental source: GEO-associated Telencephalon	Db	427 ---NPRTSFLASNAE-----NKPPVAQQSPPVLIPTLPIQTTIOTS
A;Note: sequence extracted from NCBI backbone (NCBIN:98891, NCBIP:9892)	Oy	467 C;Superfamily: protein kinase C zinc-binding repeat homology <K22>
C;Superfamily: protein kinase C zinc-binding repeat homology	Db	468 REVAPPSSINSRNAAASPFRPTISVSPOPSSPTRSLSIFGARDIAILREHSNTPNIVMQCT
F;46-95/Domain: protein kinase C zinc-binding repeat homology <K22>	Oy	527 Query Match 9 9%; Score 320; DB 2; Length 299; Best Local Similarity 31.6%; Pred. No. 4.4e-13; Matches 92; Conservative 31; Mismatches 106; Indels 46; Gaps 10; Residues 1-299
Best Local Similarity 31.6%; Pred. No. 4.4e-13; Matches 92; Conservative 31; Mismatches 106; Indels 46; Gaps 10; Residues 1-299	Db	528 259 STLNRSQLEPRTEIDSVGTPQSNGMRLHDFVSKTVTKPESCVPCCRKIRFGKLS-LKC 316
259 STLNRSQLEPRTEIDSVGTPQSNGMRLHDFVSKTVTKPESCVPCCRKIRFGKLS-LKC 316	Oy	31 ATLKENEHVPRYE-----KVNHNKVHTPRGFPHCEYCA-NFMGLIAGVKC 76
31 ATLKENEHVPRYE-----KVNHNKVHTPRGFPHCEYCA-NFMGLIAGVKC 76	Db	343 -----EVKIGEGMLADVFOSNS-----PMPSIVHCV 370
317 RDCRVVSHPECRDRCLPCTPL-----IGSPVKTGEGLMADFVFSOTSPMPISIV 366	Oy	425 REVAPPSSINSRNAAASPFRPTISVSPOPSSPTRSLSIFGARDIAILREHSNTPNIVMQCT
317 RDCRVVSHPECRDRCLPCTPL-----IGSPVKTGEGLMADFVFSOTSPMPISIV 366	Db	528 485 Q-SPHTMKVANLAKVFGPTIVAHAVENPDPPVMSODIK 522
485 Q-SPHTMKVANLAKVFGPTIVAHAVENPDPPVMSODIK 522	Oy	585 FFRELREPLIPDNHQRFIDFAGNEVEDSRRAVHRAINDLPDANTISTIRLTHLAKIK 644
77 ADCGLAVNHVKOCSKMVPNDCKPDLKHVKVVSCLDTLVK-----AHFTKR---PMVV 125	Db	645 ENDVNMKSTNNLAIWGPTLIRQAT-IPELSSFSRTE 682
77 ADCGLAVNHVKOCSKMVPNDCKPDLKHVKVVSCLDTLVK-----AHFTKR---PMVV 125	Oy	RESULT 10 T3954 probable gtpase activating protein homolog - fission yeast (Schizosaccharomyces pombe) C;Species: Schizosaccharomyces pombe C;Accession: T3954 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
R;Xiang, Z.; Aves, S.; Wood, V.; Rajanaream, M.A.; Barrell, B.G.	Db	425 FLRNKPLPFLFRNRMFAEAITDEDSNAAVQGELFQANRDTLAFLMIHQRA 484
A;Reference number: T21893	Oy	186 YFRDGLIPLITYDAFKFIESAKTIDPDEQIEILHEALKLIPPARTELRYLMHAKRV- 244
A;Accession: T3954	Db	425 485 Q-SPHTMKVANLAKVFGPTIVAHAVENPDPPVMSODIK 522
C;Status: preliminary; translated from GB/EMBL/DDBJ	Oy	528 485 Q-SPHTMKVANLAKVFGPTIVAHAVENPDPPVMSODIK 522
A;Residues: 1-695 <XAA> A;Cross-references: EMBL:AL035065; PIDN:CAA22624_1; GSPPDB:GN00067; SPDB:SPBC23G7_08C	Db	528 485 Q-SPHTMKVANLAKVFGPTIVAHAVENPDPPVMSODIK 522
A;Cross-references: strain 972h-; cosmid c23G7	Oy	585 FFRELREPLIPDNHQRFIDFAGNEVEDSRRAVHRAINDLPDANTISTIRLTHLAKIK 644
C;Genetics: A;RefSeq: SPBC23G7_08C	Db	645 ENDVNMKSTNNLAIWGPTLIRQAT-IPELSSFSRTE 682
A;Map position: 2	Oy	RESULT 11 T3954 submitted to the EMBL Data Library, January 1999 C;Species: Homo sapiens (man) C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002 C;Accession: E59430 C;Status: Aspenstrom, P.; Hellman, U.; Gomez, L.J.; Heldin, C.-H.; Saras, J.; Franzen, P.; submitted to GenBank, December 1997 A;Description: Homo sapiens PTPL1-associated RhogAP 1 (PARG1), mRNA. A;Reference number: E59430 A;Accession: E59430 A;Species: Homo sapiens (man) A;Molecule type: mRNA A;RefSeq: 1-1261 <SAR> A;Cross-references: GB:NP_004806; PIDN:NP_004806.1
Query Match 9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Db	425 RNLFEQYVRRVETLSEGNEVFIOLAKDFEDERKKMORTDHE-LGKVQDLMKAETERSA 67
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Oy	483 NRHLNQSOPSPGPGPANSLEDVRLPDSNNKTEEDRSNSAIDT--GPSFRSWTGMS 539
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Db	372 RALEEEALQKVE--EADELYKV-CYVNEVERNDVENTKREILAQRLTYFQCDITKA 427
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Oy	68 LDYKLKHARNQDVETKRRQAEADCEKLERKROIOLTRMELIPTREMIDTSQSIOLSEEQK--SA 124
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Db	428 VTVNLFLRNMHQLOAASLADRKLQASLGCSAKLYDPQGESEFVKATNS---TEEEKVDGVN 482
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Oy	125 LAFLNRQGPSSSNAGNKRISTI--DESGSILSDISDKTIESLDMDSSLVTKFLKRE 181
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Db	483 NRHLNQSOPSPGPGPANSLEDVRLPDSNNKTEEDRSNSAIDT--GPSFRSWTGMS 539
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Oy	182 KARSTSQFVODPPGPVKTTSIGSAYDQGQNSIVAKTTVWPDNGCPTEAVSTIENVY 241
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Db	540 DSEST-----GGSESRSIDS-----EST-----SPGD-----
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Oy	242 WRSRRKTGTGTOPWNSNSTLNSROLEPRTE--DSVGPQS--NGGMRLHDVSKVTK 296
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Db	563 FRRKPLPPRPSCTMSSADDLREPSPSEPGPSNLSGTFKTKLMSKALTRK--RKLRS 620
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Oy	297 PESCVPGKRKIRFGKLSLKICDCRVSHPECRDRCLP---IP---TLIGTPVKEG 349
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Db	621 PTKCRODGIVWTF--QGVECSEBLVYCHKRLENLYTCGHQKPCKG 670
9 5%; Score 308; DB 2; Length 695; Best Local Similarity 23.7%; Pred. No. 8.3e-12; Matches 137; Conservative 87; Mismatches 223; Indels 132; Gaps 23; Residues 1-695 <XAA>	Oy	350 MLADF--VSQISP-MIPSIVVHCVNTEQRSIETGGLYRISGCDRTVKELEKFLRKVTV 406

QY 199 KKTRSTGSAVDQGNESIVAKUTVTPNDDGGPIEAVSTIETVP-----YMRSRRTGTL 252
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 116 KRFESIHDLVTGLITLVIETKAA-----EYISKMTNPYHIGATLREKV--- 164

QY 253 QPNWNSDSTLNRSRQLPR-----TETUSVGTVTQSNGEM-----RLHFVSKT 293
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 165 -----SRLLSRSKSNPFRKTNTVTHEBTAVKISSILVRRAALTHNDNHNFYTKTHNPVKVHT 219

Db 294 VIKPSPSCVPGCKKRIFGKLS-LKRCRDCRVSHPSQCRDRCPLPCPTLIGTPVKGEGLM 351
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 220 FRGPFWCEYCA-NFWMGLAQGVRSQDCGIVNWKQCSKHYPNDCOPDLKRIK-KYCCDL 277
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 352 ADFVSQTSMPMPSIVHCVNBLIEQRLTETGLYRSGCDRUVKELEKEKFLR-VKIVPPL 409
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 278 TTLVKAHTNTPMVTDICIREAAGLKSEGGLYRSGFTHEIDYKMAFDRGERADISA 337
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 410 SKVDFIHAACSLKLKPFLRNKEPLTFLRRAEMRAELTDEDSNAAAMYQAVGELPQAN 469
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 338 NVYPVINITWGALKFLHFRDLPVPTVYDTSKFLAKISNADELEAVHEVMLLPPAH 397
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 470 RDTIAFLMLHQVRQAQSPHK-MDVAHLAKVFGPPTVAHVNPNNPVPVMSODIKPQKV 528
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 398 YETLYLMHLKKVYNEKMFNAENLGTVFGPTLM--RPPESTDLLTLDMDRQKLIV 455
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 529 ERLSLPLEWSQFNAMEQENI 550
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 Db 456 Q-----LLIENEDV 464
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15

D59435 Gem-interacting protein [imported] - human
 C;species: Homo sapiens (man)
 C;date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 19-Jul-2002
 C;accession: D59435
 R;Aresta, S.; Berger, R.; de Gunzburg, J.
 R;description: GMIP, a Gem interacting protein.
 A;reference number: D59435
 A;Accession: D59435
 A;status: preliminary
 A;Molecule type: DNA
 A;residues: 1-970 <ARE>
 A;Cross-references: GB:NP_057657; PID:97706107; PID:NP_057657.1

Query Match
 9.0%; Score 291; DB 2; Length 970;
 Best Local Similarity 22.6%; Pred. No. 1.6e-10;
 Matches 188; Conservative 112; Mismatches 273; Indels 260; Gaps 40;
 Cutoff 13.0; Score 291; DB 2; Length 970;
 Matches 188; Conservative 112; Mismatches 273; Indels 260; Gaps 40;
 QY 13 EQLVRLVERITLSENEGNEVQFOLOAKDFEDFRK--W-QRTDHFLGKYKDLIMKATERTSA 67
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 81 EELDLRL-IRTKGGDALEYAKTWSRYAKELLAWTEKRSAYLEFFAKSTMKIAERGKVS 139
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 68 LDVK-----LKH-----ARQVQDV----- 82
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 140 IQQSHMPLOYIYTFLFHDLSLGLTAMETVAQQRDYKOPLAARTETERKWRKFKEQW 199
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 83 IKRRGRRAEADCEKLER-QFLQR--EMLMDTSGSQSQSLREQKSALAFNLRGQQSSSNAG 139
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 200 MKEDRNMNEVAQRAQOYQVOSEDLRASQDPAQS-----BPSKQ 250
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 140 NKRJLSTIDESGSIILDISFDKTDESLLWDSSVLKV-----KRRSTS-RQFV-D 192
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 251 ERRRSRREEQA-----KAQEAEALQACVREANARQDLEAKQRIYSHVRKLVFQ 302
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 QY 193 GPPGHVKKTRSI----GSAVDQGNESIVAKUTVTPNDDG------PIE 231
 |: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 303 GDEVLRLRVTLISLFGIRGAQAAERGPRAFAALAECCAPFEGPQRYQEFVRALPAPP 362

 232 AVSTIETVPTWTRS-----RRKTCGLQ-----PRNSTD-----LNSRQEPRT 271
 Db 363 AFSPQEFPLSLNSSLDILRKLSGLPDRDENSEAPGPMEPDGPGWQHGTPGTPG 422



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:44:41 ; Search time 25 Seconds

Scoring table: BLOSUM62

Sequence: US-09-881-736-2

Perfect score: 3243

Score: 1 MDTMMINVRNIEQLVRRVE.....SKSATNLGRCQNNFFASPMK 632

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Swissprot: 40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query	Length	DB ID	Description
1	477	14.7	384	1 RN_DROME	P40809_drosophila
2	348	10.7	1980	1 MY9B_RAT	063358_rattus_noir
3	339.5	10.5	2114	1 MY9B_MOUSE	099y05_mus_musculus
4	329	10.1	334	1 CHIN_RAT	P30337_rattus_noir
5	328	10.1	334	1 CHIN_MOUSE	091v57_mus_musculus
6	323.5	10.0	2158	1 MY9B_HUMAN	Q13459_homo_sapiens
7	322	9.9	459	1 CHIN_HUMAN	P15882_homo_sapiens
8	294	9.1	295	1 CHIO_RAT	P03070_rattus_noir
9	293.5	9.1	468	1 CHIO_HUMAN	P52757_homo_sapiens
10	290.5	9.0	814	1 OPHL_HUMAN	Q9uhal_homo_sapiens
11	282	8.7	859	1 ABR_HUMAN	Q12979_homo_sapiens
12	280.5	8.6	747	1 YN54_CAEEL	P34588_caenorhabditis_elegans
13	275.5	8.5	1271	1 BCR_HUMAN	P11274_homo_sapiens
14	260.5	8.0	802	1 OPHL_HUMAN	060690_homo_sapiens
15	251.5	7.8	986	1 RHG6_MOUSE	054834_mus_musculus
16	244	7.5	802	1 OPHL_MOUSE	Q99j31_mus_musculus
17	238	7.3	974	1 RHG6_HUMAN	Q43182_homo_sapiens
18	232	7.2	666	1 RGDL YEAST	P38339_saccharomyces_cerevisiae
19	228.5	7.0	837	1 YLES_CAEEL	P46941_caenorhabditis_elegans
20	206.5	6.4	1275	1 YA99_SCHPO	Q10164_schizosaccharomyces_pombe
21	202.5	6.2	946	1 RHG4_HUMAN	P98171_homo_sapiens
22	197.5	6.1	638	1 RHO53_HUMAN	P42331_homo_sapiens
23	192	5.9	718	1 RGAL_HUMAN	Q9nsg0_homo_sapiens
24	191	5.9	425	1 RHBG_MOUSE	Q9cxp4_mus_musculus
25	188.5	5.8	439	1 RHG1_HUMAN	P07960_homo_sapiens
26	187	5.8	2167	1 BEM2_YEAST	P39960_saccharomyces_cerevisiae
27	180	5.6	622	1 RBP1_HUMAN	P9y313_homo_sapiens
28	171	5.3	650	1 Y411_HUMAN	Q43295_homo_sapiens
29	169.5	5.2	1007	1 RGAL_YEAST	P39083_saccharomyces_cerevisiae
30	169	5.2	601	1 3BPL_MOUSE	P55194_mus_musculus
31	160.5	4.9	1017	1 LRG1_YEAST	P35688_saccharomyces_cerevisiae
32	159.5	4.9	1128	1 BEML3_YEAST	P32873_saccharomyces_cerevisiae
33	142	4.4	728	1 P85B_HUMAN	Q00459_homo_sapiens

RESULTS

RESULT 1

ID RN_DROME STANDARD; PRT: 384 AA.

AC P40809; Q9V91;

DT 01-FEB-1995 (Rel. 31, created)

DT 01-FEB-1995 (Rel. 31, last sequence update)

DT 15-JUN-2002 (Rel. 41, last annotation update)

DE Grtase activating protein rracGAP.

GN RNRACGAP OR CG2395.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydriidea; Drosophilidae; Drosophila.

OX NCBL_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC STRAIN=Oregon-R; TISSUE=Pupae;

RX MEDLINE=9302458; PubMed=1406685;

RA Apel M., Roeder L., Vola C., Griffin-Shea R.;

RT "A Drosophila rotund transcript expressed during spermatogenesis and imaginal disc morphogenesis encodes a protein which is similar to human Rac Gtase-activating (racGAP) proteins.";

RT Mol. Cell. Biol. 12:5111-5122(1992).

RN [2]

RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.

RX MEDLINE=96194886; PubMed=8654933;

RA Hoellmann C.D., Bergeret E., Guichard A., Griffin-Shea R.;

RT "Alternative splicing of the Drosophila melanogaster rotund racGAP gene.", Gene. 168:135-141(1996).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC STRAIN=Berkeley;

RX MEDLINE=2019006; PubMed=10731132;

RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoxkins R.A., Gallo R.F., George R.A., Lewis S.E., Richards M., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazquez R.G., Champe M., Preteifer B.D., Wan K.-H., Doyle Y.-H., Helt G., Nelson C.R., Miklos G.L.G., Abril J.-F., Agbayani A., An H.-J., Andrews Pfrankovich C., Baldwin D., Balliew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Besson K.Y., Benos P.V., Berneman B.P., Bhanderi D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Buttier H., Cadieu E., Centter A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin R.J., Evangelista C.C., Ferraz C., Farriera S., Fleischmann W., Rosler C., Gabrieleian A.E., Gaarg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harrin N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

ALIGNMENTS

RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merklikov G., Mishina N.V., Mobley C., Morris J., Mosherf A., Mount S.M., Moy H., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svartas R., Tector C., Turner R., Venter C., Wang A.H., Wang X., Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang J., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., RA	"The genome sequence of <i>Drosophila melanogaster</i> ."; Science 287: 2185-2195 (2000).	Db
RA	-1- FUNCTION: INVOLVED IN THE MORPHOGENESIS OF THE ADULT APPENDAGES. -1- GTPASE-ACTIVATING PROTEIN FOR P21-RAC. PROMOTES THE EXCHANGE OF RAC-BOUND GDP BY GTP.	Qy	
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING.	Db	
CC	-1- TISSUE SPECIFICITY: IN PUPAE, EXPRESSED IN IMAGINAL DISKS AND ONLY IN THE MALE GONAD. IN ADULTS, ONLY FOUND IN THE TESTES, IN REGIONS FILLED WITH PRIMARY SPERMATOCYTES.	Qy	
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED DURING SPERMATOGENESIS, IN PRIMARY SPERMATOCYTES, AND IMAGINAL DISK MORPHOGENESIS.	Db	
CC	-1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG BINDING DOMAIN.	Qy	
CC	-1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.	Db	
CC	-1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-23 IS THE INITIATOR.	Qy	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	Db	
RP	MYB RAT	RESULT 2	
ID	MYB RAT	MYB RAT	
AC	063350;	STANDARD;	
DT	16-OCT-2001 (Rel. 40, Created)	PRT;	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	1980 AA.	
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Myosin Ixo (Unconventional myosin-9b).		
GN	MYOB OR MYS5.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TAXID=10116;		
RN	[1]		
RP	"A novel type of myosin implicated in signalling by rho family GTPases.",		
RC	SEQUENCE FROM N.A.		
STRAIN-Sprague-Dawley; TISSUE-Brain stem, and Spinal cord;			
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Myosin Ixo (Unconventional myosin-9b).		
GN	MYOB OR MYS5.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
RA	Baehtier M.; Reinhard J., Scheel A.A., Diekmann D., Hall A., Ruppert C.,		
RA	PIR; A48122; A48122.		
RA	FlyBase; Flygn005843; rnRACGAP.		
DR	InterPro; IPR0211; DAG_PE-bind.		
DR	InterPro; IPR00198; RhogAP.		
DR	Pfam; PF00130; DAG_PE-bind; 1.		
DR	Pfam; PF00520; RhogAP; 1.		
DR	SMART; SM00109; C1; 1.		
DR	SMART; SM00324; RhogAP; 1.		
DR	PROSITE; PS00479; DAG_PE_BIND_DOM; 1.		
DR	PROSITE; PS50081; DAG_PE_BIND_DOM; 2; 1.		
DR	Guanine-nucleotide releasing factor; Spermatogenesis; Testis; Alternative splicing; Zinc: Phorbol-ester binding.		
KW	FT PHORBOL-ESTER AND DAG BINDING.		
FT	DOMAIN 87 136 PHORBOL-ESTER AND DAG BINDING.		
FT	DOMAIN 188 202 ARG/LYS-RICH.		
FT	VARSPLIC 378 384 NLSTNL - RRMVSL (IN ISOFORM 2).		
SQ	SEQUENCE 384 AA; 44217 MW; D24959B833BT/A2AB CRC64;		
Query Match	14.7%; Score 477; DB 1; Length 394;		
Best Local Similarity	40.7%; Pred. No. 2e-23; DR		
Matches	110; Conservative 44; MisMatchs 102; DR		
DR	DR		
DR	HSSP; P08799; IMND.		
DR	InterPro; IPR00219; DAG_PE-bind.		
DR	InterPro; IPR00048; IQ_region.		
DR	InterPro; IPR000159; RA_domain.		
DR	InterPro; IPR00198; RhogAP.		
DR	InterPro; IPR001609; myosin_head.		
DR	Pfam; PF00063; myosin_head; 2.		
DR	Pfam; PF00130; DAG_PE-bind; 1.		
DR	Pfam; PF00612; IQ; 4.		
DR	Pfam; PF00520; RhogAP; 1.		
DR	Pfam; PF00788; RA; 1.		
DR	PRINTS; PRO0133; MYOSINHEAVY.		

DR Pfam: PRO0130; DAG_PE-bind; 1.
 DR Pfam: PF00612; IQ; 4.
 DR Pfam: PF00620; RhogAP; 1.
 DR Pfam: PRO0788; RA; 1.
 DR PRMTS; PRO193; MYOSINHEAVY;
 DR PRODom; PD00355; myosin_head; 2.
 DR SMART; SN00109; C1; 1.
 DR SMART; SN00115; IQ; 4.
 DR SMART; SM0242; MYSIC; 1.
 DR SMART; SM00314; RA; 1.
 DR SMART; SM00324; RhogAP; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50096; IQ; 3.
 KW Myosin; Repeat; Arm-binding; Calmodulin-binding; Actin-binding; Cytoskeleton; Coiled coil; Gpase activation; Phorbol-ester binding; Zinc; Alternative splicing; Polymorphism.

KW ZINC; Alternative splicing; Head or Motor domain.

FT DOMAIN 1 939 NECK OR REGULATORY DOMAIN.
 FT DOMAIN 940 1044 TAIL.
 FT DOMAIN 1045 2114 COILED COIL (POTENTIAL).
 FT DOMAIN 1562 1592 COILED COIL (POTENTIAL).
 FT DOMAIN 1839 1859 COILED COIL (POTENTIAL).
 FT DOMAIN 1915 1945 ACTIN-BINDING.
 FT DOMAIN 844 855 RHO-GAP.
 FT DOMAIN 957 977 ATP (POTENTIAL).
 FT DOMAIN 979 1000 R->RC(TGIDF)SERSELDVNAFEDIMAFYER (IN
 FT DOMAIN 1001 1023 ISOFORM 2).
 FT DOMAIN 1024 1053 MISSING (IN ISOFORM 3).
 FT DOMAIN 1539 1639 GPPAPALPPISPISPLPEAAAPPRGRPTSFVTVRKVPR
 FT DOMAIN 1671 1820 RPPIMPMANKLPPGLPLILISWAPALQAVVVPKRREPPA
 FT DOMAIN 239 246 RQDQHVSHTYATPGADLPQSTILALDHNLIPGSKRYSRSD
 FT VARSPLIC 740 740 PPTVCLPPPSQGOANG -> E (IN ISOFORM 3).
 FT VARSPLIC 1875 1890 L->LEVSPLVPLSSL (ONLY IN STRAIN C57BL/6; CONTAINS AN IMPERFECT DUPLICATION).

FT VARIANT 1243 1243 A -> P.

FT SEQUENCE 2083 2083 2114 AA; 238832 MW; D774DAB1B2788045 CRC64;

Query Match, Best Local Similarity 10.5%; Score 339.5; DB 1; Length 2114; Matches 167; Conservative 110; Mismatches 276; Indels 201; Gaps 28; Qy 50 ELGRKVDLIMKAETERSALDVKLKHARNQDVKEKKRQAEDDKLERQIOLR- EML 107
 Db 1244 ESPRKD--KDESTKAD--KPESPSGSTQIQRQ--HPDTERLATAVTRGKKLA 1295
 Qy 108 MCDTISGSIOUSEEOK--SALAFENRGQSSNSNAGNKRISTIDSGSISDISDK--- 160
 Db 1295 SAVLQSQSLDSEKHKATGAAITPTERRIFSTSDISKLSPVKTSAEIDQDFSSRKPSIH 1355
 Qy 161 -----TDESU-----DWQSSLVITFKLKKREKRS-----T 186
 Db 1356 KKKSQDPSAGPDAGLSQGDSKSAFKLFLHKAKDQPKSLEGVEETHSNGQQAQETP 1415
 Qy 187 SROTQVDPG-----PGVVKRPSIG---SAVGNESETVAKT-----T 220
 Db 1415 ARKTLDVPSQSQHRTGKPLKGKKNRKRKVQQTIVSKWRESVERKTINANELKFDE 1475
 Qy 221 --VTPND---GGPTEA-----VSTIETVP----- 240
 Db 1475 FLLNKVNDLRSOKPIESLPIATERFRNKTMSVPGKRIHVGYKDLMENYQIVVSNL 1535
 Qy 241 -----WYRSRRRTGTQLOPWNSTLNRSQLERETTDVSGPQ 279
 Db 1536 AAERGEKDNTNLVNVYFQSLIDEFTSYNKT-----DEFRAKSKOKKRRKOERV--Q 1586
 Qy 280 SNGGMRHLDFVSKIVKPKSCVPGKRIKGKLSLCROCRVSHPECRDRCPFLPCPT- 338

Db 1587 EHNG--HVFAFYQVNIPQSCQCSLYIWLMKALICSVKMTCHKKCVKQSYCSTG 1643
 Db 339 -----LIGTVKIGGMALAFVQSPTSPMPSPSIVHCVNFEQGDTETGILYRISGCDRTV 393
 Qy 394 KELKESKFLRVKTVPLSKVDD--IAICSLUKDFLRNKLEPLTFRLNRAFMEEAITEDE 451
 Db 1702 RELRQ---AIQTDPAVKLEDFPIHAITGVLKQNMRELEPLMFFAQYQDFRLAVELPEK 1758
 Db 1644 RRKSHGAEGHF--GVCVDSLTSKASPVIVENLLEHMGUYTEGLYRSGAANET 1701
 Qy 452 DNSIAAMYQAVGELLOQANRDTLAFMHQLRVA-QSPHRTMDVANLAKVFGPTIVAHAVP 510
 Db 1759 QEQLSRAIYAVLDHUPANEHTSLERLIFHLKVVALLEDVNRMSPGALAIIFAPCIL-RCPD 1817
 Qy 511 NPDPTYMSQDIKQPKVWEELSLPLEYNSQMMFMEQENDPLHVIEN-----SN 560
 Db 1818 NSDPDTSMSKDVYLKLTTCVMIK--EQMKRYKM-KMEELNHLEAESAFARRSLLORN 1873
 Qy 561 A-----FSTPOTPDIVKSLLGPVTPEHOLKTPSSSSLSQRVFS--- 600
 Db 1874 APWPWKLGFSSPSYEVRIKSPTVQDIELGALSSEEAGGDDEREKEIIMERIQSIKE 1933
 Qy 601 --TILTNTPRGSKRSATNLGROONFFASPMK 632
 Db 1934 KEDTYRLPELDPROSDEENLDSETSASTESLLE 1967

RESULT 4
 CHIN RAT ID CHIN RAT STANDARD: PRT: 334 AA.
 AC P30337; DT 01-APR-1993 (Rel. 25, Created)
 AC P30337; DT 01-APR-1993 (Rel. 25, Last sequence update)
 DE N-chimaerin (NC) (N-chimaerin) (Alpha chimaerin)
 GN CHN1 OR CHN.
 OS Ratmus norvegicus (Rat)
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Buteraria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=93074974; PubMed=1455199;
 RA Lim H.H., Michael G.J., Smith P., Lim L., Hall C.;
 RT "Developmental regulation and neuronal expression of the mRNA of rat n-chimaerin, a p21rac GAP/cDNA sequence.";
 RL Biochem. J. 287:415-422(1992).
 RT "- FUNCTION: GTPase activating protein for p21-rac and a phorbol-ester receptor. May play an important role in neuronal signal-transduction mechanisms."
 CC -!- TISSUE SPECIFICITY: IN NEURONS IN BRAIN REGIONS THAT ARE INVOLVED IN LEARNING AND MEMORY PROCESSES.
 CC -!- DEVELOPMENTAL STAGE: INCREASES IN AMOUNT DURING BRAIN DEVELOPMENT COINCIDENT WITH SYNAPTOGENESIS.
 CC -!- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORBOL-ESTER AND DAG BINDING DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X67250; CAM7672.1; -.
 DR PIR; S25152; S25152.
 DR PIR; S29128; S29128.
 DR HSSP; P28867; 1PQ.
 DR Interpro; IPR002219; DAG_PE-bind.

RT "Cloning of the murine unconventional myosin gene Myo9b and
RT identification of alternative splicing.";
RL Gene 24(3):389-398(1999).
RN [3]
RP SEQUENCE OF 1940-2158 FROM N.A. (LONG ISOFORM).
RX MEDLINE=98158729; PUBMED=9490038;
RA Post, P.L.; Bokoch, G.M.; Mooseker, M.S.;
RT "Human myosin-Txb is a mechanochemically active motor and a GAP for
J. Cell Sci. 111:941-950(1998).
RN [4]
RN SEQUENCE OF 1828-2023 FROM N.A. (SHORT ISOFORM).
RC TISSUE=Placenta;
RA Isogai, T.; Ota, T.; Hayashi, K.; Sugiyama, T.; Otsuki, T.; Suzuki, Y.;
RA Nishikawa, T.; Nagai, K.; Sugano, S.; Ishibashi, T.; Fujimori, K.;
RA Tanai, H.; Kimata, M.; Watabane, M.; Hiracka, S.; Ishii, S.; Kawai, Y.;
RA Saito, K.; Yamamoto, J.; Wakamatsu, A.; Nakamura, Y.; Nagahari, K.;
RA Masuno, Y.; Kaneko, K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB 2001) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
MAY BE INVOLVED IN THE REMODELING OF THE ACTIN CYOSKELETON. BINDS
ACTIN WITH HIGH AFFINITY BOTH IN THE ABSENCE AND PRESENCE OF ATP
AND ITS MECHANOCHEMICAL ACTIVITY IS INHIBITED BY CALCIUM IONS.
CC ALSO ACTS AS GTPASE ACTIVATING PROTEIN ON RHO.
-1- SUBCELLULAR LOCATION: IN UNDIFFERENTIATED CELLS COLOCALIZES WITH
F-ACTIN IN THE CELL PERIPHERY WHILE IN DIFFERENTIATED CELLS ITS
LOCALIZATION IS CYTOPLASMIC WITH THE HIGHEST LEVELS IN THE
PERINUCLEAR REGION.
-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
-1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN PERIPHERAL BLOOD
LEUKOCYTES AND AT LOWER LEVELS IN THYMUS, SPLEEN, TESTIS,
PROSTATE, OVARY, BRAIN, SMALL INTESTINE, AND LUNG.
-1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
-1- SIMILARITY: CONTAINS 1 ZINC-DEPENDENT PHORB-ESTER AND DAG
BINDING DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U42391; AAC50402.1; ALT_SEQ.
DR EMBL; AF020267; AAC25697.1; ALT_INIT.
DR EMBL; AK002201; BAQ92132.1; ALT_INIT.
DR HSSP; P08799; IMND.
DR Genew; HGENC;7609; MYO9B.
DR MM; 602129; -.
DR InterPro; IPRO02219; DAG_Pe-bind.
DR InterPro; IPRO00048; IQ_region.
DR InterPro; IPRO00159; RA_domain.
DR InterPro; IPRO00198; RhogAP.
DR InterPro; IPRO01609; myosin_head.
DR Pfam; PF00063; myosin_head; 2.
DR Pfam; PF00130; DAG_pe-bind; 1.
DR Pfam; PF00612; IQ_4.
DR Pfam; PF00620; RhogAP; 1.
DR Pfam; PF00788; RA; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR Prodom; PD000355; myosin_head; 2.
DR SMART; SM00109; C1; 1.
DR SMART; SM00015; IQ; 3.
DR SMART; SM00242; MYSC; 1.

DR SMART; SM00314; RA; 1.
DR SMART; SM00324; RhogAP; 1.
DR PROSITE; PS00479; DAG_Pe_BIND_DOM; 1.
DR PROSITE; PS50081; DAG_Pe_BIND_DOM_2; 1.
DR PROSITE; PS50056; IQ; 3.
KW Myosin; Repeat; ATP-binding; Calmodulin-binding; Actin-binding;
KW Cytoskeleton; Coiled coil; Gtpase activation; Phorbol-ester binding;
KW Zinc; Alternative splicing.
FT DOMAIN 1 939 HEAD OR MOTOR DOMAIN.
FT DOMAIN 1 940 1044 NECK OR REGULATORY DOMAIN.
FT DOMAIN 1 1045 2158 COILED COIL (POTENTIAL).
FT DOMAIN 1 1046 1071 COILED COIL (POTENTIAL).
FT DOMAIN 1 1880 1901 COILED COIL (POTENTIAL).
FT DOMAIN 1 1959 1989 COILED COIL (POTENTIAL).
FT DOMAIN 1 844 855 ACTIN-BINDING.
FT DOMAIN 1 957 977 IQ 1.
FT DOMAIN 1 979 1000 IQ 2.
FT DOMAIN 1 1001 1023 IQ 3.
FT DOMAIN 1 1024 1053 IQ 4.
FT DOMAIN 1 1633 1681 PHORBOL-ESTER AND DAG BINDING.
FT DOMAIN 1 1713 1862 RHO-RAP.
FT NP_BIND 23.9 246 ATP (POTENTIAL).
FT VARSPlice 2022 2023 PP -> QY (IN SHORT ISOFORM).
FT VARSPlice 2024 2158 MISSING (IN SHORT ISOFORM).
FT CONFLICT 1937 1939 QYP -> KT (IN REF. 4).
FT CONFLICT 1947 1947 L -> P (IN REF. 4).
FT CONFLICT 2040 2045 TVAPP -> PWPPH (IN REF. 3).
FT CONFLICT 2049 2049 P -> L (IN REF. 3).
FT CONFLICT 2067 2067 P -> S (IN REF. 3).
FT CONFLICT 2157 2158 NG -> MAEHS (IN REF. 3).
SQ SEQUENCE 2158 AA; 243556 MW; 4978F1D77056D28 CR664;

Query Match Best Local Similarity 10.0%; Score 323.5; DB 1; Length 2158;
Matches 149; Conservative 117; Mismatches 284; Indels 119; Gaps 27;

QY 22 LSEGNEVQ-----FIOALKDFDFERKWMW--RTDHDLGKYKDLMKAETEASLDV 70.
Db 1404 LSPGSQDSKSTFKFLHLTKTD---KKYSLGAEELLENAVGHVLLAETMKGLEA 1458

QY 71 ---KUKHARQVDWEIKRQRQAEDCEKLERQIQLIREMLMCDSGSIOLSECKSALA 125

Db 1459 PSGQDHRHAGE-----KRKKEPGKGKKNRNVKI-----GKTVSEKWRSEVF 1502

QY 127 FLNRGQPSSSNAGNKLRLSTIDE-SGSILSDISFDFTDESIDWDSSSLVKPFKLKREKRRS 185

Db 1503 ---ROTNTAN---ELKVYDEFLINKINDLSQTP-----IESLFEATEKFRS 1545

QY 186 TSRQFVGDPPPKVKTCSAVDGQNSAVTAKTWTVPNDGGPIEAWSITP-VYWT 244

Db 1546 NIKTMYSVPNGKI---HVGYKDLMENYQIVSNILATERQKDDNLVLNLFQSLDEFTR 1601

QY 245 SRRKGUTLQPKNNSDSTSNSRQEPLPTEDSVGTPSNSNGGVRLLHPFVSKVTKPSCVPG 304

Db 1602 GYTK-----NDFEPVKQSKAQKKRKEROAVOEHNG---HYASYQVSIPQSCOECL 1650

QY 305 KRIKFGKLSQLKCRDGRVSHSPECRCPRLPCIPY--IGTP-VKIGE-GMLADVSQISP 360

Db 1651 SYIWMDKALCSVQKMTCKVKVQIKOSHCSVTGKRGKPGEGHGFVGVDSLTSKRA 1710

QY 361 MIPSTVHCNEIEGQGLTEGLYRISGCDRTVKLKEFLRVKVPPLSKVDD---THAI 418

Db 1711 SVPVILEKLLHEVHMGLYTEGLYRISGCDRTVKLKEFLRVKVPPLSKVDD---THAI 1767

QY 419 CSLKDFLRNUKEPLITERNRAFEEATDEDENSIAAMYOAVGELPOANRDTLAFMI 478

Db 1768 TGVLKWLRLPELPPIMLTAQYGDFLRAVELPEKGQBLAAIYAVLHLPEARNHLERIF 1827

QY 479 HLQRYA-QSPHTKMDQVANLAKVGETIVAHAVPNPDVPMQSQDKRQPKVERUJSLPL 537

Db 1828 HLVKALLEDVNRMSPGALAFPCLL-RCPDNDSPLTSMKDVKITCVCMEMLIKEQM 1886

QY 538 YWSQSMNVEGENIDPLHVIEN-----SNA-----ESTPQTPDIKVSLGPVTT 580

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR Genew; HGNC:1944; CHN2.

DR MIM; 602857; -.1;

DR InterPro; IPR002219; DAG PE-bind.

DR InterPro; IPR00198; RhoGAP.

DR Pfam; PF0017; SH2; 1.

DR Pfam; PF00130; DAG PE-bind; 1.

DR Pfam; PF00620; RhoGAP; 1.

DR PRINTS; PR00008; DAGPEDOMAIN.

DR PRODOM; PD000093; SH2; 1.

DR SMART; SM00109; CL; 1.

DR SMART; SM00324; RhoGAP; 1.

DR SMART; SM0255; SH2; 1.

DR PROSITE; PS00479; DAG PE_BIND_DOM; 1.

DR PROSITE; PS00081; DAG PE_BIND_DOM_2; 1.

DR PROSITE; PS50001; SH2; 1.

KW GTPase activation; Phorbol-ester binding; zinc; SH2 domain;

FT DOMAIN 59 127 SH2.

FT DOMAIN 215 264 PHORBOL-ESTER AND DAG BINDING.

FT DOMAIN 291 435 RHO-GAP.

FT CONFLICT 1 5 MAASNN->MRLL (IN REF. 1).

SEQ SEQUENCE 468 AA; 53923 MW; 63254958E0B5804C CRC64;

Query Match 9.1%; Score 293.5; DB 1; Length 468;

Best Local Similarity 25.7%; Pred. No. 1.4e-11;

Matches 98; Conservative 66; Mismatches 63; Gaps 12;

Indels 15;

Qy 199 KKTRSIGAVDQGNPSIVAKTVTVFNDGGPIEAVSTIETPV----WVTRSRKTGT 252

Db 118 KRFESIHDLYTDGLJTLVYETKAA-----EVISKMTPNPIVEHIGATLLRKV-- 166

Oy 253 QPWNSDSTLNSRQLPFR-----TEIDSVGTPOSNGM-----RLHDIFVSKT 293

Db 167 ----SRRLSRSRNKEPRKTNTTHEHTAVAKISSLVRRAALTHNDNHNFYEKTHNFKVHT 221

Qy 294 VIKPESCVPGKRIKGKLS--LKRCRCSVSHPCDRDRCPCLPCTPLIGTPVKIGEML 351

Db 222 FRGPWHCEYCA--NFWMGLAQVGRSDGGLNVHKRCSKRVNDCPDLDKRIK-KYTCQL 279

Oy 352 ADFVQSOTSPMIPSTIVWCNEORGLTETGLYRTSGCORTVKELKEKFLR--VKTVPILL 409

Db 280 TTLVKAHTQPRPMVMDICIREBANGKSKSGLYRVSFGTFEHEDKMAFDRGKADISA 339

Oy 410 SKVDDITHAICSLSLKUFLRNKEPLITFLRMRAFMEAETDEDSNIAAMQAVGSLPPQAN 469

Db 340 NVPDPDINTGALKLYFDRDLPPIPVTYDKSFDAKISNADERLEAHHEVMLLPPAH 399

Oy 470 RDTLAFLMILHORVQPHTK-MDVANLARVFGPTIVAHAVPNPDPTVMSQDIKQPKV 528

Db 400 YETLRIMLMIKKVVMNEKNFMNAENLGIVFGPTLM--RPPESTDSTLTLHDMDRYQKLIV 457

Oy 529 ERLLSLPLEWSQFMVMEQEINI 550

Db 458 Q-----ILINEDVY 466

RESULT 10
 OPHL_HUMAN STANDARD; PRT; 814 AA.
 ID OPHL_HUMAN
 ID O9UNAI; 075117; Q9UJ00; Q9BY56; Q9BY57;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Oligophrenin-1 like protein (GTPase regulator associated with focal adhesion kinase).

DE OPHNL OR GRAF OR KIAA0221.

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI-TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RP TISSUE=Placenta, Heart, and Liver;

RA Xia J.-H., Tang X.X., Yu K.P., Pan Q., Dai H.P.;

RT "Molecular cloning of human oligophrenin-1 like (OPHNL) gene, complete CDS"; Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2), DISEASE, AND VARIANT LEUKEMIA SER-417.

RX MEDLINE=20381355; PubMed=10908648;

RA Borkhardt A., Bojesen S., Haas O.A., Fuchs U., Bartelheimer D., Lencarevic T.R., Bohle R.M., Harbott J., Repp R., Jaeger U., Viehmann S., Henn T., Korth P., Scharr D., Lampert F.;

RT "The human GRAF gene is fused to ML in a unique t(5;11)(q31;q23) and both alleles are disrupted in three cases of myelodysplastic syndrome/acute myeloid leukemia with a deletion 5q.";

RA [3]

PROC. NATL. ACAD. SCI. U.S.A. 97:9168-9173(2000).

RN [4]

RP SEQUENCE OF 53-785 FROM N.A. (ISOFORMS 1 AND 2).

RA Bojesen S.E., Link C., Borkhardt A.;

RT "Genomic structure of the human GRAF gene.";

RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.

RP SEQUENCE OF 62-814 FROM N.A. (ISOFORM 1).

RC TISSUE=Brain;

RX MEDLINE=9840380; PubMed=974811;

RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.;"

RT DNA Res. 5:165-176(1998);

CC FUNCTION: GTPase activating protein for RhoA.

CC -I- SUBUNIT: Binds to the C-terminal of pp25(FAK).

CC -I- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.

CC -I- DISEASE: A form of Juvenile myelomonocytic leukemia is characterized by a chromosomal translocation t(5;11)(q31;q23) that involves OPHNL and ML.

CC -I- SIMILARITY: Contains 1 PH DOMAIN.

CC -I- SIMILARITY: Contains 1 RHO-GAP DOMAIN.

CC -I- SIMILARITY: Contains 1 SH3 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; AF41184; AAD39482.1; .

DR EMBL; Y10388; CAY71414.2; .

DR EMBL; AJ09466; CAC29145.2; .

DR EMBL; AJ09467; CAC29145.2; JOINED.

DR EMBL; AJ09468; CAC29145.2; JOINED.

DR EMBL; AJ30946; CAC29145.2; JOINED.

DR EMBL; AJ30940; CAC29145.2; JOINED.

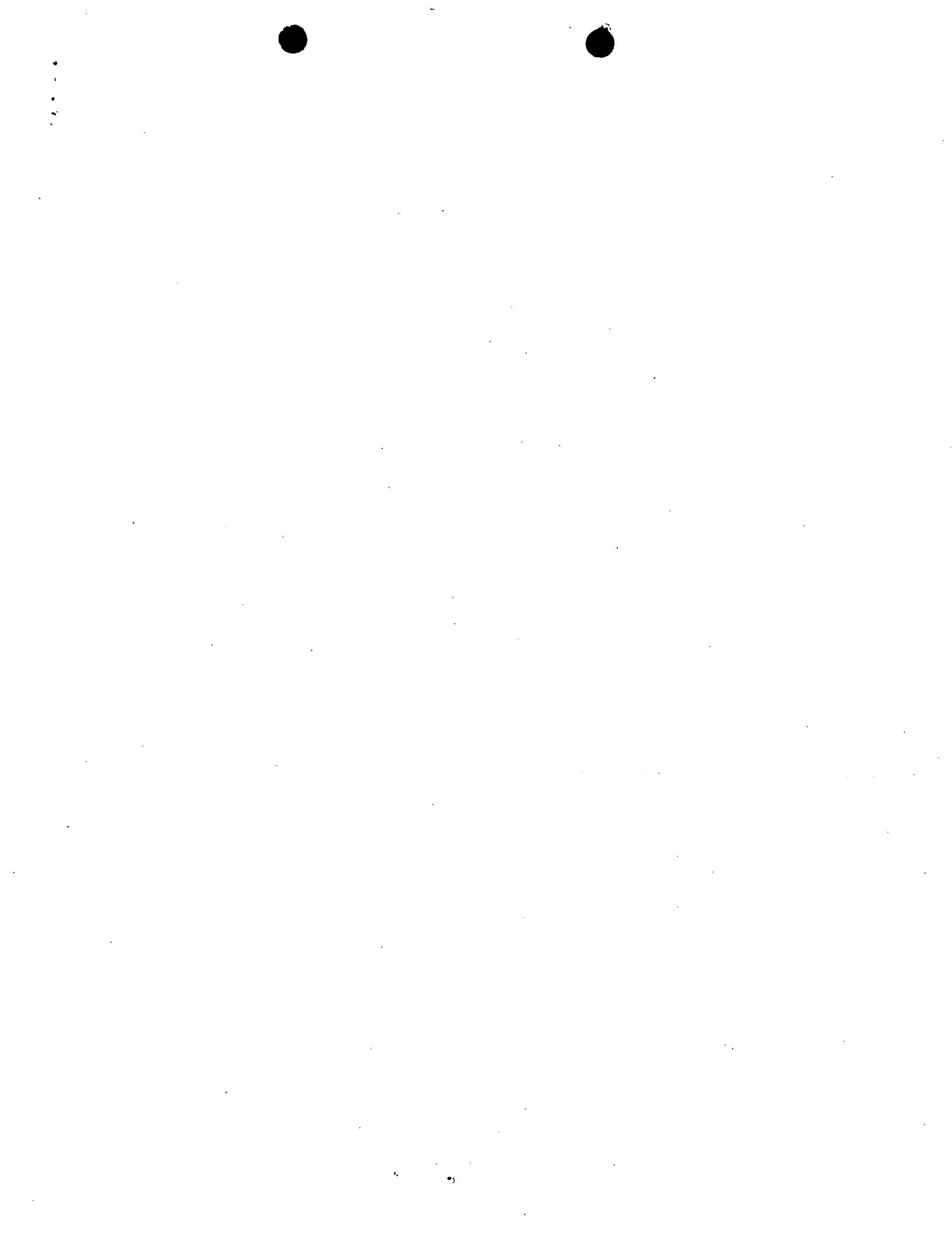
DR EMBL; AJ30947; CAC29145.2; JOINED.

	Matches	137; Conservative	86; Mismatches	196;	Indels	153;	Gaps	28;
CC	BCR-ABL ONCOGENE.							
CC	- I - PTM: AUTOPHOSPHORYLATED.							
CC	- I - DISEASE: Participates in a t(9;22)(q34;q11) chromosomal translocation that produces a BCR-ABL oncogene responsible for chronic myeloid leukemia (CML), acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL).							
CC	- I - SIMILARITY: CONTAINS 1 DBL_HOMOLOGY (DH) DOMAIN.							
CC	- I - SIMILARITY: CONTAINS 1 PH DOMAIN.							
CC	- I - SIMILARITY: CONTAINS 1 C2 DOMAIN.							
CC	- I - SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.							
CC	- I - SIMILARITY: STRONG_TO HUMAN ABR AND DROSOPHILA ROTUND.							
CC	This SWISS PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.ebi-sib.ch/announce/or_send_an_email_to_licensee@ebi-sib.ch).							
CC	EMBL; U07000; AAB60388_1; -							
DR	EMBL; M24603; AAB65939_1; -							
DR	EMBL; Y00661; CAH6876_1; -							
DR	EMBL; X02596; CAZ2644_1; -							
DR	EMBL; M15025; AAA35594_1; -							
DR	EMBL; X52828; CAH3701_1; -							
DR	EMBL; M64337; ; NOT_ANNOTATED_CDS.							
PIR	A26172; TVHUB2.							
PIR	A28765; TVHUR3.							
DR	Genew; HGNC:1014; BCR.							
MIM	151410; -							
DR	InterPro; IPR000008; C2.							
DR	InterPro; IPR001331; GDS_CDC24.							
DR	InterPro; IPR001849; PH.							
DR	InterPro; IPR000219; RHOGAP.							
DR	Pfam; PF00168; C2; 1.							
DR	Pfam; PF00159; PH; 1.							
DR	Pfam; PF00620; RHOGAP; 1.							
DR	Pfam; PF00621; RHOGEF; 1.							
DR	SMART; SM00239; C2; 1.							
DR	SMART; SM00233; PH; 1.							
DR	SMART; SM00224; RHOGAP; 1.							
DR	PROSITE; PS50004; C2_DOMAIN_2; 1.							
DR	PROSITE; PS50003; PH_DOMAIN; 1.							
DR	PROSITE; PS00741; DH; 1; 1.							
DR	PROSINE; PSS0010; DH; 2; 1.							
KW	Guanine-nucleotide releasing factor; Proto-oncogene; Transferase; Chromosomal translocation; Serine/threonine-protein kinase; Phosphorylation.							
FT	DOMAIN 1 426 KINASE.							
FT	DOMAIN 498 691 DH.							
FT	DOMAIN 708 866 PH.							
FT	DOMAIN 870 1002 C2 DOMAIN.							
FT	DOMAIN 1068 1212 RHOGAP.							
FT	BINDING 197 385 TO ABL SH2-DOMAIN.							
FT	DOMAIN 824 877 POLY-LGL.							
FT	SITE 426 427 BREAKPOINT FOR TRANSLOCATION TO FORM BCR-ABL ONCOGENE.							
FT	CONFFLICT 287 287 M -> T (IN REF. 4).							
FT	CONFFLICT 418 418 G -> D (IN REF. 4).							
FT	CONFFLICT 483 483 E -> K (IN REF. 4).							
FT	CONFFLICT 560 560 F -> S (IN REF. 4).							
FT	CONFFLICT 733 733 E -> D (IN REF. 1 AND 4).							
FT	CONFFLICT 796 796 S -> N (IN REF. 1).							
FT	CONFFLICT 961 1004 MISSING (IN REF. 4).							
SQ	SEQUENCE 1271 AA; 142805 MW; A59F76B#1AF3236 CRC64;							
Query Match	8.5%; Score 275.5; DB 1; Length 1271;							
Best Local Similarity	24.0%; Pred. No. 7.5e-10;							

Query Match Score 275.5; DB 1; Length 1271;
Best Local Similarity 24.0%; Pred. No. 7.5e-10;

Matches 137; Conservative 86; Mismatches 196; Indels 153; Gaps 28;
QY 43 KW-----QRDHELGKYDLMKAKERTSADLVKHKARNQDVKEIKRRQRAEADC 93
Db 754 KWYVPLTDLSFOMD-ELEAVPNIPVLPDEELDAALKI---SOIKSDIOPRKANGS 808
QY 94 EKLERQIQLIREMMACDTSGSIQSEEOKSALAFFLNRGQPSSNAGNRSLTIDESGIL 153
Db 809 KATER---LKKL-----SEGBSLLMSPSMAFRVHSRNGKSYF---LIS 849
Db 154 SDISFDKTDIDESLDWSSLUVKTEKLKKRKRSTSQRQFYDGPGRPVKKRSISAVDQ-G 211
Db 850 SDY-----ERAERENI-----REQOKCFRSF-----SLISVELQMLT 883
Db 944 KAKRVYVYDVTPLAEPWNEFIELEGESQPLRILYEVCKYNKTIPKEDGESTDRMLGKGQV 1003
Db 884 NSCVKLQTVPHSIPLTINKEDDESPLGFLNVTVHSATGFKQSSNLNTYLEVDSFGFVN 943
QY 243 -TRSRRKGITLQP-WNSDSL--NSROLE-----PTREDSVGTPQSNGM 284
Db 336 IPTLIGTPVKGECMLAPVFSQSPMIPSIVWICNEIQRGLTETGLYRISCDRTWKE 395
Db 1048 QTGVFG--VVI-----AVVTKREKSKVPUYVROCEEETERRGMEBVGYRVSGVATDQA 1100
QY 396 LKEKF-LRKVTPVPLJSKVDIHAICSLSLKDFLRNLIKEPLTFRLNRAEMAETIDEDNS 454
Db 1101 LKAFDGVNKKDVSQSMSEMDNAAGTUKYFLPBLPEFTDFEYPAEGLASDPAV 1160
QY 455 IAAMYQAVGELPOQNRTLAFLMTHLQRYAQSPH-TKDMDVANLAKVQFPTIV----AH 507
Db 1161 ESCMLNLISLPEANLTFPLFLDHLKRVAAEAVNKMSLHNLTAVFOPTLRSERESK 1220
QY 508 AVPPRP-DPVTM---SODIKRKPWVERULS 534
Db 1221 LPANPSQPTMTDSSWSLEVMSQVQVQLYFLQL 1252

RESULT 14
OPHI_HUMAN
ID OPHI_HUMAN STANDARD; PRM; 802 AA.
AC 068890; Q8WXK7; DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Created)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN O14gphrenin 1.
OPHL
OS Homo sapiens (Human).
OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Bivalvia; Eutherrinia; Primates; Cattarrini; Homino.
OC NCBI_TAXID=9006;
[1] RNP
RC SEQUENCE FROM N.A., AND DISEASE.
RC TISSUE-Fetal brain;
RX MEDLINE-98241228; PubMed=9820272;
RA Billuart P., Bienvenu T., Roncet V., Vinet M.C.,
RA Zemni R., Roest Crollius H., Caillie A., Fauchereau F., Cherry M.,
RA Brillaut S., Hamel B., Fryns J.-P., Beldjord C., Kahn A., Moraine C.,
RA Cheilly J.;
RT "Oligophrenin-1 encodes a rhogap protein involved in x-linked mental retardation,"
RT Nature 392:923-926(1998).
RN [2]
RN SEQUENCE FROM N.A., AND VARIANTS THR-45 AND MET-301.
RX PubMed=10812214;
RA Billuart P., Cheilly J., Carrie A., Vinet M.C., Couvert P.,
RA McDonell N., Zemni R., Kahn A., Moraine C., Beldjord C., Bienvenu T.,
RA "determination of the gene structure of human oligophrenin-1 and



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2003, 09:56:22 ; Search time 55 Seconds
(without alignments)

1321.373 Million cell updates/sec

Title: US-09-881-736-2

Perfect score: 3243

Sequence: 1 MDTMMLNVRNLFEQLVRRVE.....SKSATNLGROGNFFASPMLK 632

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 440863 seqs, 114992915 residues

Total number of hits satisfying chosen parameters: 440863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications _AA:*

```

1: /cgn2_6_ptodata/1/pubbaa/US08_NEW_PUB.pep;*
2: /cgn2_6_ptodata/1/pubbaa/PCT_NEW_PUB.pep;*
3: /cgn2_6_ptodata/1/pubbaa/US06_PUBCOMB.pep;*
4: /cgn2_5_ptodata/1/pubbaa/US06_PUBCOMB.pep;*
5: /cgn2_6_ptodata/1/pubbaa/US07_PUBCOMB.pep;*
6: /cgn2_6_ptodata/1/pubbaa/US07_PUBCOMB.pep;*
7: /cgn2_6_ptodata/1/pubbaa/PCTUS_PUBCOMB.pep;*
8: /cgn2_6_ptodata/1/pubbaa/US08_PUBCOMB.pep;*
9: /cgn2_6_ptodata/1/pubbaa/US09_NEW_PUB.pep;*
10: /cgn2_6_ptodata/1/pubbaa/US09_PUBCOMB.pep;*
11: /cgn2_6_ptodata/1/pubbaa/US10_NEW_PUB.pep;*
12: /cgn2_6_ptodata/1/pubbaa/US10_PUBCOMB.pep;*
13: /cgn2_6_ptodata/1/pubbaa/US60_NEW_PUB.pep;*
14: /cgn2_6_ptodata/1/pubbaa/US60_PUBCOMB.pep;*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	3243	100.0	632	10 US-09-881-736-2
2	3230	99.6	632	10 US-09-833-790-413
3	2119	90.0	570	10 US-09-833-790-427
4	2734	84.3	628	10 US-09-881-736-4
5	541.5	16.7	681	10 US-09-881-736-6
6	294.5	9.1	193	10 US-09-802-127-8
7	286.5	8.8	186	9 US-10-153-668-212
8	285.5	8.8	1354	9 US-10-153-668-470
9	270.5	8.3	248	10 US-09-851-682A-1
10	265.5	8.1	555	9 US-09-764-865-878
11	251	7.7	170	9 US-10-080-960-32
12	251	7.7	170	10 US-09-802-127-7
13	241.5	7.6	291	9 US-09-764-868-898
14	223.5	6.9	152	9 US-10-152-585-5
15	218	6.7	103	9 US-10-080-960-34
16	213.5	6.4	429	9 US-09-764-868-1029
17	210.5	6.5	803	10 US-09-760-698A-2
18	206.5	6.4	257	9 US-09-764-868-1024
19	6.3	881	10	US-09-816-860A-2

Database : Listing first 45 summaries

ALIGNMENTS

RESULT 1
Sequence 1, Appl
Sequence 2, Application US/09881736
; Sequence 2, Application US/09881736
; Patent No. US20020076785A1
; GENERAL INFORMATION:
; APPLICANT: Glotter, Michael
; APPLICANT: Jantsch-Plunger, Verena
; APPLICANT: Romano, Alper
; APPLICANT: Misima, Masanori
; APPLICANT: Kaitna, Susanne
; TITLE OF INVENTION: CYK-4 polypeptides, DNA molecules encoding them and their use.
; TITLE OR INVENTION: Screening methods
; FILE REFERENCE: 0632_2260001/ERKAS
; CURRENT APPLICATION NUMBER: US/09/881,736
; CURRENT FILING DATE: 2001-06-19
; PRIORITY APPLICATION NUMBER: EP-00 112 880.0
; PRIORITY FILING DATE: 2000-06-19
; PRIORITY APPLICATION NUMBER: EP 01 110 554.1
; PRIORITY FILING DATE: 2001-04-30
; PRIORITY APPLICATION NUMBER: 60/241,231
; PRIORITY APPLICATION NUMBER: To be determined
; PRIORITY FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 2
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-881-736-2
Query Match Score 100.0%; Score 3243; DB 10; Length 632;
Best Local Similarity 100.0%; Pred. No. 2.7e-239;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 OKSALAFLNQGPSSSNAGNKRSLTIDESGSILSISDFDTDESLWDSSLVKTFKLKKR 180
 Db 121 OKSALAFLNQGPSSSNAGNKRSLTIDESGSILSISDFDTDESLWDSSLVKTFKLKKR 180
 QY 181 EKRSTSROFVGDPGPVKTRSTSGSAVDQGNESIVAKTTVTVDGGIEAVSTIETVP 240
 Db 181 EKRSTSROFVGDPGPVKTRSTSGSAVDQGNESIVAKTTVTVDGGIEAVSTIETVP 240
 QY 241 YWTSSRKGTLOPWNSDTLNSLQLEPRTEGTYGTPVKGEGMLADFSQTS 300
 Db 241 YWTSSRKGTLOPWNSDTLNSLQLEPRTEGTYGTPVKGEGMLADFSQTS 300
 Db 241 YWTSSRKGTLOPWNSDTLNSLQLEPRTEGTYGTPVKGEGMLADFSQTS 300
 QY 301 VPGCKRIKGKLSKCRDRVSPECRCPCLCPIPLIGTPVKGEGMLADFSQTS 360
 Db 301 VPGCKRIKGKLSKCRDRVSPECRCPCLCPIPLIGTPVKGEGMLADFSQTS 360
 QY 361 MIPSTVWCNEIFORGLTETGLYRISGDRTVKELEKFLRVKTVPLSKVDDIHACS 420
 Db 361 MIPSTVWCNEIFORGLTETGLYRISGDRTVKELEKFLRVKTVPLSKVDDIHACS 420
 QY 421 LKODFLRNKEPLIFRLNRAFMAETTEDNSIAAMQAVGELPQANRDTLAFMLH 480
 Db 421 LKODFLRNKEPLIFRLNRAFMAETTEDNSIAAMQAVGELPQANRDTLAFMLH 480
 Db 421 LKODFLRNKEPLIFRLNRAFMAETTEDNSIAAMQAVGELPQANRDTLAFMLH 480
 QY 481 ORVAQSPHTKMDVNLAKYFGPTVIAHAVPNPDPTVMSODIKRQPVVERLISPLEWS 540
 Db 481 ORVAQSPHTKMDVNLAKYFGPTVIAHAVPNPDPTVMSODIKRQPVVERLISPLEWS 540
 QY 541 QFMVEQEIDPLHVIENNSNAFSQTQTPDIKVSLGLGPVTPEROLKIPSSSLQRVRS 600
 Db 541 QFMVEQEIDPLHVIENNSNAFSQTQTPDIKVSLGLGPVTPEROLKIPSSSLQRVRS 600
 QY 601 TLTKNPTRGSKSATNLGRQGFASPMLK 632
 Db 601 TLTKNPTRGSKSATNLGRQGFASPMLK 632
 QY 601 TLTKNPTRGSKSATNLGRQGFASPMLK 632
 Db 601 TLTKNPTRGSKSATNLGRQGFASPMLK 632

RESULT 2
 US-09-833-790-413
 ; Sequence 413, Application US/09833790
 ; Patent No. US20020062288A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Sechrist, Heather
 ; APPLICANT: Mohamath, Radoh
 ; APPLICANT: Indrias, Carol Y.
 ; APPLICANT: Fan, Liqun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.512
 CURRENT APPLICATION NUMBER: US/09/833,790
 CURRENT FILING DATE: 2001-04-11
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 413
 LENGTH: 632
 TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-833-790-413

Query Match 99.6%; Score 3230; DB 10; Length 632;
 Best Local Similarity 99.7%; Pred. No. 2.6e-238; Matches 630; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MDTMLNVLNLFOLVRMELTSLGNEVFIOLAKDFEPRKKWORTDHELGKDLMK 60
 Db 1 MDTMLNVLNLFOLVRMELTSLGNEVFIOLAKDFEPRKKWORTDHELGKDLMK 60
 QY 61 AETERSALVVLKHARNQDVETKRRORREADCEKLERIOQLREMLMDTSSIONSE 120
 Db 61 AETERSALVVLKHARNQDVETKRRORREADCEKLERIOQLREMLMDTSSIONSE 120

QY 63 TERRALDVLKHARNQDVETKRRORREADCEKLERIOQLREMLMDTSSIONSE 122
 Db 1 TERSALDVLKHARNQDVETKRRORREADCEKLERIOQLREMLMDTSSIONSE 120
 QY 123 SALAFLNQGPSSSNAGNKRSLTIDESGSILSISDFDTDESLWDSSLVKTFKLKKR 182
 Db 61 SALAFLNQGPSSSNAGNKRSLTIDESGSILSISDFDTDESLWDSSLVKTFKLKKR 182

RESULT 3
 US-09-833-790-427
 ; Sequence 427, Application US/09833790
 ; Patent No. US20020062288A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Sechrist, Heather
 ; APPLICANT: Mohamath, Radoh
 ; APPLICANT: Indrias, Carol Y.
 ; APPLICANT: Fan, Liqun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 FILE REFERENCE: 210121.512
 CURRENT APPLICATION NUMBER: US/09/833,790
 CURRENT FILING DATE: 2001-04-11
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 427
 LENGTH: 570
 TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-833-790-427

Query Match 90.0%; Score 2919; DB 10; Length 570;
 Best Local Similarity 99.8%; Pred. No. 1.3e-214; Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 4
 Query : TRSRKKGTLQWNSDSTLNSROLEPRTEDSVGTPOSNGGRLHDPSVSKTWIKPESCP 302
 Db : PSIVHCNEIQRGLTETGGLYRISGCDRTVBLKEFLRKVPLSKVDDIHACSSL 422
 Db : PSIVHCNEIQRGLTETGGLYRISGCDRTVBLKEFLRKVPLSKVDDIHACSSL 360
 Db : KOPFLNKLKEPLTFLRNRAFMEAETTDENDSIAAMYQAVGELPQANRDTAFLMHQR 482
 Db : KOPFLNKLKEPLTFLRNRAFMEAETTDENDSIAAMYQAVGELPQANRDTAFLMHQR 420
 Db : VASPHKMDVANLAKVFGPTIVAHAVPNPDVMTMSQDIKRPKVVERLLSPLWWSQF 542
 Db : VASPHKMDVANLAKVFGPTIVAHAVPNPDVMTMSQDIKRPKVVERLLSPLWWSQF 480
 Db : MAVEQENIDPLAVIENSNAFSPTQPDTKVSLIGPVTTPEHOLKTTPSSSLSQRVSTL 602
 Db : MAVEQENIDPLAVIENSNAFSPTQPDTKVSLIGPVTTPEHOLKTTPSSSLSQRVSTL 540
 Db : TKNTPRGSSKSATNLGROGNEFASPMILK 632
 Db : TKNTPRGSSKSATNLGROGNEFASPMILK 570

RESULT 4
 Sequence 4, Application US/09881736
 Patent No. US20020076785A1

GENERAL INFORMATION:
 APPLICANT: Glotzer, Michael
 APPLICANT: Jantsch-Plunger, Verena
 APPLICANT: Romano, Alper
 APPLICANT: Mishima, Masanori
 TITLE OF INVENTION: Cyt-4 polypeptides, DNA molecules encoding them and their use in screening methods

CURRENT APPLICATION NUMBER: US/09/881,736
 CURRENT FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: EP 00 112 880.0
 PRIOR FILING DATE: 2000-06-19
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 4
 LENGTH: 628
 TYPE: PRT
 ORGANISM: Mus musculus

Query Match 84.3%; Score 2734; DB 10; Length 628;
 Best Local Similarity 84.4%; Pred. No. 2e-200; Matches 534; Conservative 47; Mismatches 46; Indels 6; Gaps 3;

Query : MOTMMILVRNLIEQYRVEILSEGND-WOFLOAKDFEDERKKWQTHDELGKYKDLML 59
 Db : MOTMMILVRNLIEQYRVEILSEGND-WOFLOAKDFEDERKKWQTHDELGKYKDLML 60

RESULT 5
 Sequence 6, Application US/09881736
 Patent No. US20020076785A1

GENERAL INFORMATION:
 APPLICANT: Glotzer, Michael
 APPLICANT: Jantsch-Plunger, Verena
 APPLICANT: Romano, Alper
 APPLICANT: Mishima, Masanori
 TITLE OF INVENTION: Cyt-4 polypeptides, DNA molecules encoding them and their use in screening methods

CURRENT APPLICATION NUMBER: US/09/881,736
 CURRENT FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: EP 00 112 880.0
 PRIOR FILING DATE: 2000-06-19
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 6
 LENGTH: 681
 TYPE: PRT
 ORGANISM: Caenorhabditis elegans

Query Match 16.7%; Score 541.5; DB 10; Length 681;

Query Match 60 KAETERSALDVKURHARNQDVDEKERRAEADOEKELRQIQIREMAMCDTSQSILOSE 119
 Db : KARTGSRSLDVKURHARNQDVDEKERRAEACAKLEQQIQLIRDIMCDTSQSILOSE 120
 Db : EOKSALATLNSRGPSSVAGNKLUSTIDESGSTLSDISFDKTDLSLDWDSLVKTFLKK 179
 Db : RERRSTSRSQFVGPPGVKIRSSIGSAVQGENESTVAKTTVPGPTEAVSTIETVYW 180
 Db : EOKSALAFLNROGASSGHAGNNRLISTIDESGSISFDKTDLSLDWDSLVKNMK 180
 Db : RERRSTSRSQFVGPPGVKIRSSIGSAVQGENESTVAKTTVPGPTEAVSTIETV 239
 Db : REKRSNSRQFDGGPGPKCSIGSTVDOANESIVAKTTVPGPTEAVSTIETL 240
 Db : PYWTRSRSQRTGTQWNSDSTLNSROLEPRTEDSVGTPOSNGGRLHDPSVSKTWIKPES 299
 Db : PSWRSRSQKSGPQPVNSDALSNSRPLEPRTEDSVGTQNTGGMRLHDPSVSKTWIKPES 300
 Db : CVCGKTRKGKSLKCRDCRVSHSPCDRDPLCIPITLGPVKIGEMLADFVQS 359
 Db : CVCGKTRKGKSLKCRDCRVSHSPCDRDPLCIPITLGPVKIGEMLADFVQS 360
 Db : SLKDFLNLKEPLTFLRNRAFMEAETTDENDSIAAMYQAVGELPQANRDTAFLMH 479
 Db : SLKDFLNLKEPLTFLRNRAFMEAETTDENDSIAAMYQAVGELPQANRDTAFLMH 480
 Db : LORVAQSPTKMDVANLAKVFGPTIVAHAVPNPDVMTMSQDIKRPKVVERLLSPLWYSW 539
 Db : LORVAQSPTKMDVANLAKVFGPTIVAHAVPNPDVMTMSQDIKRPKVVERLLSPLWYSW 540
 Db : SPTKMPRFDTKMDIANLAKVFGPTIVAHAVPNPDVMTMSQDIKRPKVVERLLSPLWYSW 599
 Db : SPTKMPRFDTKMDIANLAKVFGPTIVAHAVPNPDVMTMSQDIKRPKVVERLLSPLWYSW 541
 Db : NQFMVMDQENDS---ORGNGNSTPRDPVDSVLLGPVTPPEFOLVKTPLSSLSORY 596
 Db : 597 NLKSTPREGNKSKSATNLGODGKFPPAYLK 628

Best local Similarity 25.9%; Pred. No. 6.4e-33; Matches 180; Conservative 109; Mismatches 285; Indels 121; Gaps 22;

QY 7 NVRNLFEODVRYRVEILSBNCENEVOFIQLARDFEDFRKKW-----ORNDHELGKYKDLM 59

Db 14 NSRHIFNMILNSQRPOFDKIDGFMHLIDEIERLKLWMDSEESKKRUMADMBRAEAEA 73

Qy 60 KAETERSALDVKLKHARNOVDEVELRKRRQFAADECEKLERQIQLTREMULMCDTSGSIQLE 119

Db 74 KARKKLAMPDIDVKTOKHILALMEENKAALKLDUNYERREOKLDAMKGNTFS--LTK 131

Qy 120 EOKSALAFITNRGQSSSNAGNKRSLIDESGSIISLSDFDKTDDESLDWD--SSLVKFF 176

Db 132 EDROQFKELHE--PLVRITYSKRVQ--QRFPHMIEDTQDEDSEVDYDDGOSFEEVH 186

Qy 177 LKK--REKRESTS-----RQFYQPPG---PVK 200

Db 187 LNRGREVRRSSAAGNAVGGKRRSASAHAATTAANSKRSRSRVMATIDEPBNEGGTPKR 246

Qy 201 TRSISGSAVQDGNESIVAKTVT-----VPNDGDP-----IEAVSTIEVPP 240

Db 247 CRDGGSTPQHQEMTPTTTTTTINHSRQNQDPDPRVSHRLQTLRRSLSCGSIPSCDQTP 306

Qy 241 YWTRSSRRKGTQLQPNNSDSTLNSROLEPPTETSVGTPOSNGGNRLHDFVSKVVIKPSSC 300

Db 307 GQTNNIGGMSSALTNSTDIRTDIRKRGTPAWNGTR--DIAMRPHFIEAGIKAMRK 365

Qy 301 VPGCKRTIKKGKLSLKCRDRVVSPECRCRCPICP-----TLLGTPVKIGCM 350

Db 365 DKCATALKA-TSMKCRDQHOVWHRSCCNKLHPCIPREKTMMPKSAORGKPGAGER 424

Qy 351 LADEFVSQTSPPMIPSIVHCVNELEORGLETGLYRISGCDRTYKELKEFLRVKTVPLS 410

Db 425 LODLCTSAPKMPIAVHCVVALEARGLJQEGIYRPGCVRTVNVLDB-LRSKTPVNG 483

Qy 411 KVDIHAICSSLKDLFRNUKEPLTFRNRAFMEAET--TDEDSIAQMVAQGEFLQA 468

Db 484 -LHDVEVTDTLKRFLDKPLIPRTSRQELTVAANLYSTDDNGRLANRVCIEQPA 542

Qy 469 NRDTLAFMLHLQRV-AQSPHTKDVANIAKVGPTVHAVNPDPVMSQ-----D 520

Db 543 NRDLAYLFLTHWRKVIAQSSRNKNCAMARMVAPAVMH-----PVQSQOAIAGR 596

Qy 521 IKROPKVVERLLSLPLEWSQMFAMEQEONIDPLHVIENSNAFSTPQTPD---IKVSLIG 576

Db 597 ATDCHRAMTALEFERDDVWQRFGLTS-----VSMASNOETARHODNFALCDRSILG 649

Qy 577 PVTPPEHOLKTSSSISQRVSITLTNTPFGS 611

Db 650 PVIT-----SPATPLLARSANATRARGAHLLGS 677

RESULT 6
S-09-802-127-8
Sequence 8, Application US/09802127
Patient No. US2002045212A1

; GENERAL INFORMATION:
; APPLICANT: Gluckmann, Maria Alexandra
; APPLICANT: Meyers, Rachel

; TITLE OF INVENTION: No. US2002045212A1 Human GTPase Activator Proteins

; FILE REFERENCE: 038001/58994 US/09/802,127

; CURRENT FILING DATE: 2001-02-23

; PRIORITY APPLICATION NUMBER: 611,185,611

; PRIORITY FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 193

; TYPE: PRT

; ORGANISM: Artificial Sequence

; OTHER INFORMATION: Rho-Gap3 consensus sequence

US-09-802-127-8
Query Match 9.1%; Score 294.5; DB 10; Length 193;

Best Local Similarity 40.1%; Pred. No. 7.5e-15; Matches 81; Conservative 26; Mismatches 54; Indels 41; Gaps 8;

Qy 359 SPMFESTIVHCVNTEORGLTETGLYRISGCDRIVKEKEFIRKVYKUPLSKVD---- 413

Db 1 SP-IFIIVEKCIEYLEKRGDLTGYRYVSGSKSRVKELEAFDSDGEDD--LDSLDESITE 57

Qy 414 -----DIHAICSSLKDLFRNUKEPLTFRNRAFMEAET-----TDE 451

Db 58 ESEDEELEYVHDWAGLKLHYRLEPELPELFEEYEFEAKLYQIEATSRKOSKESEDE 117

Qy 452 DNSIAMYOANGELIPQANUDTAFMLMIHQRYAQ-SPIHKMDVANLAKVFGPTIVAHAVP 510

Db 118 EERIRALRGLSLSLIPANERATVYL-LHNHRVKEHSEVNKMARNLAIVFGPLLR--- 172

Qy 511 NPDTVMSODIKRQPVYRLL 532

Db 173 --PPLT--DIKHONKVWETLI 189

RESULT 7
US-10-153-668-212
Sequence 212, Application US/10153668
Publication No. US20030092616A1

; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi

; APPLICANT: MATSUUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya

; TITLE OF INVENTION: STAT6 Activating Gene
; CURRENT APPLICATION NUMBER: US/10/153, 668

; PRIORITY FILING DATE: 2002-05-24
; PRIORITY APPLICATION NUMBER: US 60/293, 172

; PRIORITY FILING DATE: 2001-05-25
; PRIORITY APPLICATION NUMBER: US 60/316, 031

; PRIORITY FILING DATE: 2001-08-31
; PRIORITY APPLICATION NUMBER: US 60/328, 403

; PRIORITY FILING DATE: 2001-10-12
; PRIORITY APPLICATION NUMBER: JP 2001-157043

; PRIORITY FILING DATE: 2001-05-25
; PRIORITY APPLICATION NUMBER: JP 2001-260681

; PRIORITY FILING DATE: 2001-08-30
; PRIORITY APPLICATION NUMBER: JP 2001-313175

; PRIORITY FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488

; SEQ ID NO 212
; LENGTH: 1286

; TYPE: PRT
; ORGANISM: Homo sapiens

; US-10-153-668-212
Query Match 8.8%; Score 286.5; DB 9; Length 1286;
Best Local Similarity 22.1%; Pred. No. 4.8e-13; Matches 143; Conservative 98; Mismatches 216; Indels 191; Gaps 26;

Qy 48 DHE-LGKTK-DLMKAEFERSL-----DKVL--KHARNOVDEVIEKRRQFADC 93

Db 187 DHESVGPFLSDQPNKSTERSKSYEGDDYRDQAKLSFKHSSLKQIKADSQKSEEDS 246

Qy 94 -----EKERQOLIREMLMCDSGSIOLSEBOKS 123

Db 247 GSRKDSSESVFDAAKEGCWLHFRPLVTDKGKRVGGSRIPWKWQMVLRGHSLYKDKRE 306

Qy 124 ALAFLNRQPSNSNAGNKRSLIDESGSIISLSDFDKTD----- 161

Db 307 QTPSEEEBOP1SYNA-----CLIBDISYSEMKRKNVFRLTSDCBLFQAEADR 353

Qy 162 DESLDWDSLVKFLKKREKRSTSKOFVGDPPGVKTRSIGSAVQDGNESIVAKTV 221

RESULT 8
US-10-153-668-470
; Sequence 470, Application US/10153668
; Publication No. US20030092610A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAKA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
FILE REFERENCE: 1254-0207P
CURRENT APPLICATION NUMBER: US/10/153, 668
PRIOR APPLICATION NUMBER: US 60/293, 172
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/316, 031
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: US 60/328, 403
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: JP 2001-157043
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: JP 2001-260681
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: JP 2001-313175
PRIOR FILING DATE: 2001-10-10
NUMBER OF SEQ ID NOS: 488
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 470
LENGTH: 1354
TYPE: PRT
ORGANISM: Homo sapiens
US-10-153-668-470

Query Match 8.8%; Score 286.5; DB 9; Length 1354;
Best Local Similarity 22.1%; Pred. No. 5.2e-13; Matches 143; Conservative 98; Mismatches 216; Indels 191; Gaps 26;

Qy 48 DIB-LGYK-DILMKAEPERSAL-----DVKL--KHARNQDVWEIKRQRQAEDC 93
Db 255 DBESVGPPLSDAQPNSKERSISYDEGIDDYREDAKISFKVSSLKKGKIAQSSEDS 314

RESULT 9
US-09-851-622A-1
; Sequence 1, Application US/09851682A
; Patent No. US20020091248A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Choi, Ying
; APPLICANT: Duhl, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Lieng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MOSIN IXA AND CYCLIC NUCLEOTIDE GATED
TITLE OF INVENTION: CHANNEL-15 (CNNG-15) POLYNUCLEOTIDES, POLYPEPTIDES,
TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF
FILE REFERENCE: 2001130 442
CURRENT APPLICATION NUMBER: US/09/851, 682A
CURRENT FILING DATE: 2001-05-08
PRIOR APPLICATION NUMBER: US/09/172, 422
PRIOR FILING DATE: 1998-10-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2548
TYPE: PRT
ORGANISM: Homo sapien
US-09-851-622A-1

Query Match 8.3%; Score 270.5; DB 10; Length 2548;
Best Local Similarity 24.4%; Pred. No. 2.2e-11; Matches 130; Conservative 75; Mismatches 203; Indels 125; Gaps 24;

QY 143 LSTIDESGILSDISFDKIDESLDQDSSLVTKFLKKRKRSTSROFDGPPGVKKR 202
 Db 1 :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
 1864 LKSMDE--FLLKVN-DLDNEDESKDVTLYDVKALKERPRN-----IFSHY 1908
 QY 203 SIGSAVDQGN-----ESTVAKTVTVPNDG--EPIA-VSTIETVWTRSR 247
 Db 1 :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
 1909 SSALAMDDEKSIRYKDLYALFEQLEKTRVTLQRDSLGSPPVRWNTKVFEYMNFR 1988
 QY 248 KTGTLOPWNISDSTL---NSQLEPTRTENDSVGTPQSNGMRLHDFTVTKPESCV 302
 Db 1 :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
 1969 KT-----SDCTATKVPTERKKRKETDLV-EHNG---HIFKATQSYIPTYCEY 2015
 QY 303 CGKRKFGKLKSLKRDRCRVSHPSGRCDRPLPCPLTGTPVKIGE-----GMLADIV 355
 Db 1 :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
 2016 CSSLWIMDRASVERCKLCKKCC-----CLRTTAKCSKKDPPELSRSQFGVELSRL 2068
 QY 355 SQTSPMIPPSVWHCNEIISORGLTETGLRISGDTRVKELEKELFRVTVPLSKVDD- 414
 Db 2069 TSEDRTVPLVTEKLNINYEMHGLYTEGIRKSGSPNKRKEILRQ---GLDTDAEVNLDDY 2125
 QY 415 -THAICSLIKDFLRNUKEPLTFRINRAMEAAETDEDSIAAMYOAVGELPOANRDTL 473
 Db 2125 NIRVIASVFKQWLRLPNPLMTFELYEEFLRANGLQERKETIRGVYVSDQLSKRHLNTL 2185
 QY 474 AFLMHFLORVA-QSPHTKDVANLAKVEPTIVAHAVPDPVMSQDKRQPKVERUL 532
 Db 2186 ERLTFHLVRLTAQEDTNRUSANALIAVFPACIL-RCPDTDPLSVQDISKTTCVE-- 2241
 QY 533 SLPLYWSEPMMVQEN-----IDPHVLIENSNAFQPOTDIKUSL----- 575
 Db 2242 -----LIVVEOMNKYKARLKDISSELEFAENK-----AKTRLSLIRRSMGKGRI 2284
 QY 576 -----GPVTPHEPOLLKPKPSSSLSQVRSTLTKNTPPRGSKSKSATNLGRQ 622
 Db 2285 RRGNYPGP-SSP--VVWRIPSPVSVDSE--ETLT-----SEAMETDITEQ 2324

RESULT 10
 US-09-764-868-878
 ; Sequence 878, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et. al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; PRIORITY APPLICATION NUMBER: US/09/764,868
 ; PRIORITY FILING DATE: 2001-01-17
 ; PRIORITY FILING DATE: 2000-10-20
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 878
 ; LENGTH: 555
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-764-868-878

Query Match 7.7%; Score 251; DB 9; Length 170;
 Best Local Similarity 39.8%; Pred. No. 1.3e-11; Mismatches 52; Indels 18; Gaps 5;
 Matches 64; Conservative 27; Mismatches 52; Indels 18; Gaps 5;

QY 363 PSIVHCVNIE----ORGLTETGLRISGDTRVKELEKELFRVTVPLSKVDD- 413
 Db 1 PRIVVKCVELEKYLPLAARGLQEGIYVSGSARVKERAEADKGPDSDLESEKEW 60

QY 195 PGP-----VKKRSIGAVDQGNEISIAKTTVVPNDGPIAVSTETVWTRSR 247
 Db 52 PGPGAGDFKFERNSSGT-----MSSTEELVDPGGA----- 84

QY 248 KTGTLQPWNSDSTLNSRQLPRTEDSVGTPSPOSNGMRLHDFTVTKPESCV 302
 Db 85 -----GASAFAEQADNGMTEPLPVAVP- SGPF RHEGSKARTRRLRRLTPK 132

QY 300 CVPGKGRKFGKLSLKCRCRVSHPECUDRCPLPC-ITLIGTPVKGGMLADF--VS 356
 Db 133 CRECNSYVVF---QGAECCECLACHKKCIELTAIOCGHKLQGLQFLCQ---DFSHA 186

QY 357 QTSP-MIPSIVWCNEIISORGLTETGLRISGCDRTVKELEKELFRVTVPLSKVDDI 415

RESULT 11
 US-09-802-127-7
 ; Sequence 7, Application US/09802127
 ; Patent No. US2003004221A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria Alexandra
 ; APPLICANT: Meyers, Rachel
 ; TITLE OF INVENTION: No. US2002004521A1el Human GTPase Activator Proteins
 ; FILE REFERENCE: 035800/158994

Db 187 RSAPPGVPEFTVKVCVCEIERRALRKTGIVRNVGKTRVERLCQAFENGKELVELSOASP- 245
 QY 416 HAICSLIKDFLRNUKEPLTFRINRAMEAAETDEDSIAAMYOAVGELPOANRDTL 457
 Db 246 HDISAVNLKLTQREPLISFRLTELVLAKDSIKAEEAKAASRGDGSESEAVATA 305
 QY 458 ---MQQANGELPQANRDTIAFLMHFLORVAQ-SFHTKMDVANLAKVFCPTIVAHAVNP 512
 Db 306 LAGRILLELDLPPBNRASLQYLUHRRIVEEQDNKMTPGNLGIVFGPTLR--PRP 362

Db 363 TEATVSLSSLVLDYPRHQARVIELI 386

RESULT 12
 US-10-080-960-32
 ; Sequence 32, Application US/10080960
 ; Publication No. US20030197695A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: Glucksmann, Maria
 ; APPLICANT: Meyers, Rachel
 ; TITLE OF INVENTION: 80090, 52874, 52880, 63497, AND 33425 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES THEREOF
 ; FILE REFERENCE: 38155-2004.00
 ; CURRENT APPLICATION NUMBER: US/10/080, 960
 ; CURRENT FILING DATE: 2001-10-19
 ; PRIORITY NUMBER: US 60/242, 040
 ; PRIORITY FILING DATE: 2000-10-20
 ; PRIORITY NUMBER: US 60/242, 038
 ; PRIORITY FILING DATE: 2000-10-20
 ; PRIORITY NUMBER: US 60/241, 992
 ; PRIORITY FILING DATE: 2000-10-20
 ; PRIORITY NUMBER: US 60/242, 637
 ; PRIORITY FILING DATE: 2000-10-23
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 32
 ; LENGTH: 170
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; OTHER INFORMATION: consensus amino acid
 US-10-080-960-32

Thu Jul 3 15:01:25 2003

us-09-881-736-2.rapb

Page 8

Query Match 6.7%; Score 218; DB 9; Length 103;
Best Local Similarity 45.1%; Pred. No. 2.1e-09; Matches 46; Conservative 21; Mismatches 29; Indels 6; Gaps 3;
Matches 46; Conservative 21; Mismatches 29; Indels 6; Gaps 3;

Qy 409 LSKVDDIHAIACSLIKPLRNKKEPLIPLRNLRAFMKA-AEITDEDSIAA--MQAVG 463
Db 2 MEEYEDYHTVAGLLKQYFRELPEPLITYELVEETAAKAOVSDEDERMEALEMULIK 61

Qy 464 ELPQANHDTLAFMLTHIQRVAQ-SPHMKMDVANLAKVFGPTI 504
Db 62 LLPEANRETLYKLKHLRSVAQHSEENKMAQNLAVFGPIL 103

Search completed: July 3, 2003, 10:05:54
Job time : 57 secs

OM protein - protein search, using sw model
Run on: July 3, 2003, 09:50:57 ; Search time 27 Seconds
Copyright (c) 1993 - 2003 Compugen Ltd.

SUMMARIES

OM protein - protein search, using sw model
Run on: July 3, 2003, 09:50:57 ; Search time 27 Seconds
Copyright (c) 1993 - 2003 Compugen Ltd.

ALIGNMENT

OM protein - protein search, using sw model
Run on: July 3, 2003, 09:50:57 ; Search time 27 Seconds
Copyright (c) 1993 - 2003 Compugen Ltd.

Db 621 PTKCRDCEGIVWF-QGVECEBCLVCHRKCLELNVLVICHQKLGKHLFG----- 670 QY 356 SQTSPMPSIVHVCNEIEORGATETGLYRISGCDRTVKELKEFLRVKTVPLSKVDD- 414 Db 2069 TSEDRTVPLVKEKLINYEMHGLTVEGIRKGSTNKIELRQ--GLDTDAESWNDDY 2125 QY 350 MLADF--VSOTSP-MIPSTIVHVCNEIEORGATETGLYRISGCDRTVKELKEFLRVKTV 406 Db 671 --AEPFLVAKRKEPDGSIPTFLKICASEIENRACLOGIYRG-----NKKTE 716 QY 407 PLLSKVD----DI---HAICSLKLDRNLNEKEPLTRLNRAFMARAE---ITDE 451 Db 415 -HAICSLKLDRNLNEKEPLTRLNRAFMARAE---ITDE 451 QY 717 KLCLALENGHMLVLDSEFFSHDICOLVLYTROLPEPFLFRYKEFIDAKETOHVNEE 776 Db 452 DNS-----IAWYQAVGELRQANRDTLAFMLMIHQVA-OSPHR 490 QY 777 QETKKNLSERDKWPNCMCIEINRLLIKSKDLRQLPASNFSNLHILVHKLKVDAEENK 836 QY 491 MDVANLAKVPGPTIVAHAVNPND-PVTMSQ-DIKRQPKVERULSPLLEYWSQFM-- 543 Db 837 MNSKNLGIVFGPSLR--PRPQTAPISTSLSLAESNQARLVEFLT---YSOKIFDG 888 QY 544 -MVEQENTIDPLHVIENS--NAFPSTPQTPDVKVSLGPVTPEROLIKTPSSSSLQRVR 599 Db 889 SLQPODVMSIGVWQPGCPRKPLLSPEERDIERSMKSLEFSSKEDIHTSESEKIFERAT 948 QY 600 STLTKNTPRIGSKSISATNLGR 621 Db 949 S-----FEESERKQNALGK 952

RESULT 2
US-09-172-422-1

Sequence 1, Application US/09172422A

GENERAL INFORMATION:

- PATENT NO. 630085
- APPLICANT: Adams, Arven E.
- APPLICANT: Chiu, Choi Ying
- APPLICANT: Duhl, David
- APPLICANT: Gorman, Susan W.
- APPLICANT: Leng, Song
- APPLICANT: Sheffield, Val
- APPLICANT: Welch, Juliet

TITLE OF INVENTION: MYOSIN IIA AND CYCLIC NUCLEOTIDE GATED POLYPEPTIDES, POLYNUCLEOTIDES, COMPOSITIONS, METHODS, AND USES THEREOF

TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES, CURRENT APPLICATION NUMBER: US/09/172,422A

CURRENT FILING DATE: 1998-10-14

NUMBER OF SEQ ID NOS: 3

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1 LENGTH: 2548 TYPE: PRT ORGANISM: Homo sapien

6-09-172-422-1

Query Match 8.3%; Score 270.5; DB 4; Length 2548;
Best Local Similarity 24.4%; Pred. No. 1e-15; Matches 130; Conservative 75; Mismatches 203; Indels 125; Gaps 24;

QY 143 LSTIBERSGSTLSDISFDKTDTSLSQDSSLYKTEKLKRREKRSSTSRSQFVDPGGPPVKKR 202 Db 1864 LRSMDE--FLKKRN-DDNEDSKDFTLVDFVFKALKERON-----IFSYF 1908 QY 203 SIGAVDQGN-----ESTIVAKTTVVPNNG--GPIEA-VSTTETVPTWRSRR 247 Db 1909 SSALAMDDGRSIRYKDYALFEQOLEKTHLQEQRSLGSPPRVWVNTRKVELBEYMF 1968 QY 248 KTGTLQWPNSDSTL--NSROLPERTEDSVGTPQSGMRLJDFVSKTVPKESCP 302 Db 1969 KT-----SDCTATKVPTERKKRKETDLV--EEHNG--HIFKATQSYIPTYCE 2015 QY 303 CGKRKFGKSLSLKCDRCVYSHPCDRDPLCPIPLCIPVKIGE-----GMLADRV 355 Db 2016 CSSLIWIMDRASVCKLCKVACHKKC-----CLRTTAKCSKKYDPPELSRSQFGVELSRL 2068

RESULT 3
PCT-US93-03076-10

Sequence 10, Application PCT/US9303076

GENERAL INFORMATION:

- APPLICANT: Whitehead Institute for Biomedical Research
- TITLE OF INVENTION: GAM-Associated Protein p190 and Title of Invention: Transduction
- NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

- ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
- STREET: 2 Militia Drive
- CITY: Lexington
- STATE: MA
- COUNTRY: US
- ZIP: 02173

COMPUTER READABLE FORM:

- MEDIUM TYPE: floppy disk
- COMPUTER: IBM PC compatible
- OPERATING SYSTEM: PC-DOS/MS-DOS
- SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

- APPLICATION NUMBER: PCT/US93/03076
- FILING DATE: 19930331
- CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

- NAME: Granahan, Patricia
- REGISTRATION NUMBER: 32,227
- REFERENCE/DOCKET NUMBER: WHI92-03A
- TELECOMMUNICATION INFORMATION:
- TELEPHONE: 617-861-6240
- TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

- LENGTH: 165 amino acids
- TYPE: AMINO ACID
- TOPOLOGY: linear
- MOLCULE TYPE: protein

PCT-US93-03076-10

Query Match 7.9%; Score 257.5; DB 5; Length 165;
Best Local Similarity 37.9%; Pred. No. 1.8e-16; Matches 64; Conservative 33; Mismatches 63; Indels 9; Gaps 4;

QY 369 CVNEEQRGITETGLYRISGCDRTVKELKEFLR--VKTVPPLSKVDDTAICSLKLGF 426 Db 1 CIREFSEGRGINSSEGIVYRVSGFSDLVEDVMAFDGEGAKDISNMVENDINITGALKYF 60 QY 427 RNLKPLPLTFRNRFMEAETTDEDSNAAWYQAVGELRQANRDTLAFMLMIHQVRQS 486 Db 61 RDLPPLPLTFRNRFMEAETTDEDSNAAWYQAVGELRQANRDTLAFMLMIHQVRQS 486

RESULT 4 PCT-US93-03076-9
 Sequence 9, Application PC/TUS9303076
 GENERAL INFORMATION:
 APPLICANT: Whitehead Institute for Biomedical Research
 TITLE OF INVENTION: GAP-Associated Protein p190 and
 TITLE OF INVENTION: Transduction
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: 2 Militia Drive
 CITY: Lexington
 STATE: MA
 COUNTRY: US
 ZIP: 02113
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/03076
 FILING DATE: 19930331
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: WHI92-03A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 140 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 PCT-US93-03076-2
 Query Match 7.1%; Score 230; DB 5; Length 140;
 Best Local Similarity 38.8%; Pred. No. 5, 4e-14;
 Matches 54; Conservative 24; Mismatches 59; Indels 2; Gaps 2;
 QY ||| 369 CVNEIEQRLTETGLYLYNSGCDRTVKELKEFK-LRKVTPVLISKVDDITHAICSLKDFLFR 427
 Db 1 CVEEIERGMRGMEVIVYRSGVATDQDQLKAADFVNKNQVSVMSEMDYNAALAGTLKLYFR 60
 QY ||| 428 NKEPLITFLRNRAFMMAETTDEDDNSTAAMYQAVGELPQANRDTLAFMLHIIQRVAQSP 487
 Db 61 ELPPEPLTDEFYPNFAEGIALSPVAKESCMNLILLSLPEANILTFLELDHKRRAKE 120
 QY ||| 488 H-TKMDVYANLAKVFGPTV 505
 Db 121 AVNKMMLNHLNLTVEGPILL 139
 RESULT 5 PCT-US93-03076-2
 Sequence 2, Application PC/TUS9303076
 GENERAL INFORMATION:
 APPLICANT: Whitehead Institute for Biomedical Research
 TITLE OF INVENTION: GAP-Associated Protein p190 and
 TITLE OF INVENTION: Transduction
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: 2 Militia Drive
 CITY: Lexington
 STATE: MA
 COUNTRY: US
 ZIP: 02113
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/03076
 FILING DATE: 19930331
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: WHI92-03A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1513 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 PCT-US93-03076-2
 Query Match 6.5%; Score 210; DB 5; Length 1513;
 Best Local Similarity 19.1%; Pred. No. 2, 3e-10;
 Matches 125; Conservative 99; Mismatches 224; Indels 208; Gaps 27;
 QY ||| 67 ALDVVKLKHARNOVQDVKEVKRRDAAE-----CEKLERQIOIRELML----- 107
 Db 892 KIDV-----LNDLSREQLTEGEETIAQEDGRFTSIPOSHPOHLELFHFFKDVWE 943
 QY ||| 108 -----MCDTGSIQLSEQEKSALAFLNQNRGPSSNAGNKRISTIDESG 151
 Db 944 KNNTIIBATHMVDNVAACSTTEEVENSPRACSPPLCNSL-QDSEEDVEPPSYHLFRBDAT 1002
 Qy 152 IISDISFDKTDESLDMD----SSLVTKFLKKREKRS-TSQFDVGPPGVK----- 200
 Db 1003 -LPSLSDKHSKFSEMELEGNDGSLFSIMSNFNEKSFLNK-----VPPPKPKPVH 1050
 Qy 201 --TRSGSAVODGNESIVAKTVT-----VPNDG-----GPIEAV-----STIERYV 241
 Db 1051 ETKDL-SYLQDGHRGREGORKSNSSPPWMPQDFDPDSYAEPMDAVVKPRNEENIYSPV 1109
 Qy 242 WTRSRRTGTGTPAWN-----SDSTLNRSQLEERTEDSVGTQFQSNGGMLHDVS 291
 Db 1110 -DSTPOCKIITIRNINKAQNSNGNSGSDSEMDSLSSLERGRKVA-----VS 1153
 Qy 292 KIVIKBESCVPVGK---RIFK-----GKLSKKD----- 318
 Db 1154 KPVLYKTRCTIGRFASYRTTSFVGSDDELGIRKKEQDASQGYKGDNVIPYTEDP 1213
 Qy 319 -----CRVSHBCDRCLPLCT-----PTLIGTPVKIGEMGLADSPWIPQSOTSPM 363
 Db 1214 RRRNIILSRLRNTKPKPK-----PRPSITKAWESNYFGVP-----LTWVTPERK-IP 1262
 Qy 364 SIVVHCVNEIQRLTETGLXLRISGCDRTVKELKEKLRVVKVPLSKVDDITHAICSLK 423
 Db 1263 IFIERCIEYIATGLSTEGIYRVSGNKSEMISLQROFDDHNDLAKDKFTVTNAGMK 1322
 Qy 424 DFLRNKEPLITFLRNRAFMMAETTDEDDNSTAAMYQAVGELPQANRDTLAFMLHIIQRV 483
 Db 1323 SFFSELDPPLPYSMQLDLMVEAKINDREQKHLAKVELKPKKENHEVKVYISHLRV 1382
 Qy 484 KOSPHRK-MDVANLAKVFGPTVYAHAVPNPDPVMTSMODIKRQPKVVERLLS----- 533
 Db 1383 SHNNKVNLMTSNCWPTLMRDPDFSSMDALTATRSYQ---TILEFIQCPFFYNR 1439
 Qy 534 -----LPLEYNSQFMMVEQENIDPLHINTNSAFSTQPTPDFKVSLGPV 579

Db 1440 PISEPPGAALAPIQPHPL----SPSSPHILL-----PVSHLPSHULQPLS 1482

RESULT 6 ; Sequence 6, Application US/09415522A

US-09-415-522-6 ; Sequence 6, Application US/09415522A

GENERAL INFORMATION:

APPLICANT: Gaffney, Thomas

APPLICANT: Wendland, Juergen

APPLICANT: Philipsen, Peter

TITLE OF INVENTION: Development

FILE REFERENCE: CGC2046

CURRENT APPLICATION NUMBER: US/09/415,522A

CURRENT FILING DATE: 1999-10-08

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 6

LENGTH: 2071

TYPE: PRT

ORGANISM: Ashbya gossypii

US-09-415-522-6

Query Match 5.9%; Score 191; DB 4; Length 2071;

Best Local Similarity 22.4%; Pred. No. 2.4e-08;

Matches 128; Conservative 71; Mismatches 203; Indels 170; Gaps 25;

QY 54 YKDILMKAAEERSALDVK----LKHARNQDVEIKRRQRAEA--DCEKLERO-1Q 101

QY 102 LIREMLMCDSGSIQILOSEQKSALAFNLRGQPSNNAGNKRNLSTIDE--SGSIL---S 154

Db 1554 YDLIPFNTEQULPASHDEKAHOFGSLFLHYGTTESSIIKAERKASKEAASNEARKLFQAMG 1613

QY 1614 LFNDILVTEVY-KVQRDKQEQL-----TVOEHAEAKRSVLQHPN 1653

QY 155 DISFDKTDSEIDWDLSLVKTFKLKKREKRKRSTSROFDGPGPVKTRSIG---SAVQ 210

Db 1654 KVSVVSASSSSVSGSSS-----GSTAR-TSNPAHAAYALNMGSLSIASAHRH 1698

QY 211 GNESTIVAKT-----TVTVPNDGGPIERAVSTIETVPMWRSRRTGTQPNWSUTLNS 263

Db 1699 GRSSYSSRSSVISNTATATSPASGASPQNTISTH--HGGMKGKIGFLRPFISGFTS 1755

QY 264 RQLERTERTSVGTQNSGGM-----RLHDVFVKVIKP--ESCVP----- 302

Db 1756 SSSOTTITTSVLSGVQANGSISPYELPELSEIQTIVKIVITEIKSCIQINNYRQP 1815

QY 303 ---GGRKIRFG--KLSKRD-----CRYVSHPECRDRCPCLPCIPLTG 342

Db 1816 DMMHGFKIWMEDGIOHTQLCMDADHMENWMAKITSKRSFSHK--REFGKTSNKIFGV 1872

QY 343 PVKIGEGMLADFVSTSPIMPSIVHCVNIEORGLTETGLYRISGCDDRVKEKEFER 402

Db 1873 PVE-----DVCEREGALIPNITVKKLDTIELRGLDEVLYRPGSVSINALNAFD 1925

QY 403 VKTVPFLSKYD----IHACSLIKDFRLNLKEBLTFLNRAMEAETDEDSIAA 457

Db 1926 EGAVHNTFTLEDDRWFEEINTAGCFKLYRELPEIFT-----NEKVDEFVNINT 1975

QY 458 MYQ-----ANGELPOANRDTLAFMLHORVAQS-PHTKMDVANLAKV- 500

Db 1976 AYKNEHVDLISQFQNGIKTLISTLPFNYHNIKRFHLHNLRVHQHVNNRMADASLNAVIS 2035

QY 501 -----GPT-----TVAHAVRPD 513

Db 2036 MSFINODDIALASTMGPTLGLQMLQHLLRNPE 2067

RESULT 7

PCT-US93-03076-8

; Sequence 8, Application US/09415522A

US-09-415-522-8 ; Sequence 8, Application US/09415522A

GENERAL INFORMATION:

APPLICANT: Gaffney, Thomas

APPLICANT: Wendland, Juergen

APPLICANT: Philipsen, Peter

TITLE OF INVENTION: No. 6291660al Fungal Genes Required For No. 6291660al Growth

FILE REFERENCE: CGC2046

CURRENT APPLICATION NUMBER: US/09/415,522A

CURRENT FILING DATE: 1999-10-08

NUMBER OF SEQ ID NOS: 28

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 8

LENGTH: 1013

TYPE: PRT

ORGANISM: Ashbya gossypii

US-09-415-522-8

Query Match 5.0%; Score 161; DB 4; Length 1013;
 Best Local Similarity 21.6%; Pred. No. 5.1e-06;
 Matches 83; Conservative 61; Mismatches 128; Indels 112; Gaps 15;

Qy 178 KKRKEKRSRSRQFVGDGGG---PVKKRSIGSAVDQ-----GNRSIVARTT 220
 Db 648 KERQWISVLTLCDGPQSGTAATPSINSKSEASSLFEQTTSISDSSYLGPIANLEADMAT 707

Qy 221 VTPVNPGCPIEAVENTIETVPMWRSRRTGTQWP-----NSDSLNSROLE 267
 Db 708 PTRENDPNPVLSEEK---EVKRRMKSFFPFKKLATPTPYAACNDNASIFSODD 763

Qy 268 -PRTEDSVGTPOSNGGMRHLDEVSKVTKPECPVCPCKSKRKIKLSKCRDC-RVVSHP 325
 Db 754 SPYNATNSGGSISHSIQLSOSNLAQDYNAV-----FG---ADLSCIQLQSSH 805

Qy 326 ECRDRCPPLCIPPTLIGTPVKIGEMLADFLVSQTPMISIVVHCNET-EQRGLETG 384
 Db 806 -----YQSYEILSVWFKTLEFLYKNRQIQQEGIF 835

Qy 385 RISGCDCRYKVELKEKEFLYKVPLISKD-----DIAICSLKDFERN 428
 Db 836 RUSGSSSIKLSQEQFDKEYDVDCNYKDVSVTPGNENQGGLYVDVWVNSGLKLKLRK 895

Qy 429 LKSPBLTRLNRMAEIR---DEDSIAAMQAVGELPOANRDTLAFMLHQLRVA 484
 Db 896 L-PHMIG-DAYMDFERIVERNGDDSKLIALEFRALVNSGRIAKEVALMFAELIV 952

Qy 485 Q---SPRTKMDVANLAKVFGPTI 504

Db 953 KITENSKNMNURNLCTVFSPL 975

RESULT 9
 US-09-735-934A-2
 ; Sequence 2, Application US/09735934A
 ; Patent No. 6372468
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, Jiaxin et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEARIC
 ; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CLO00851
 ; CURRENT APPLICATION NUMBER: US/09/735,934A
 ; CURRENT FILING DATE: 2000-12-14
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSEQ For Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 878
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-735-934A-2

Query Match 4.0%; Score 128.5; DB 4; Length 878;
 Best Local Similarity 23.0%; Pred. No. 0.0048; Matches 88; Conservatve 35; Mismatches 121; Indels 139; Gaps 18;

Qy 14 QLVRVVELSEGNEVOFIOLAK-DFEDRIKMQRTDHIGKYDLIMKAERTSALDWKL 72
 Db 107 QLVRSQGQIQLQEGDULVEVVLASASMTFED---QIRPHALTWS----- 145

Qy 73 KHARNOVIVEIKRORADECEKL-ERQIQLREMIMQDTSGSIQLSEOKSALAFLNRG 131
 Db 146 -----YRAFAFCDHGEMFGFLYRQGLKDGC---LWHRKRCASFIP- 186

Qy 132 QPSSSNAGNKRKLSTIDESGSLSDISFR-KTDESDLWDWSSLVLTIFKLUKREKRSTSQF 190
 Db 187 -NCGAKRRLS---STSLASGCHSYVLGTSBSSLPCA-----EELSRTIELL 230

Qy 191 VDPPPGPKKTRLIGSATDQGNISIVATPTVTPNPGCPIEAVENTIETVPMWRSRRTGT 250
 Db 231 PRREP-----SSSSSSSSSYT---GPIEL----- 253

Qy 251 TLQWNNSDSTLNSTNQLEPRTEDSVGTPOSNGGMRHLDEVSKVTKPECPVCPCKSKRKIKLSKCRDC-RVVSHP 309
 Db 254 -----ADLSCIQLQSSH-----KPLHSTPVCQACKLKL 288

Qy 310 GKLSLICKRCRVCVSHPECRDRCPCLCI-PTLIGTPVKIGEMLADF-----VSQT 358
 Db 289 FROGLQCKDOCKFNCHKRCATRVNPDCPDLGALINGDVPMEA-TDFSEADKSALMDESED 346

Qy 359 SPMPSI---VWCYNELEQRG 377
 Db 347 SGVTPGSHSENALHASEEEBEGEG 369

RESULT 10
 US-08-655-871-2
 ; Sequence 2, Application US/0868585871
 ; PATENT NO. 6013499
 ; GENERAL INFORMATION:
 ; APPLICANT: MARUMITA, Shuh
 ; TITLE OF INVENTION: RHO TARGET PROTEIN KINASE P160
 ; NUMBER OF SEQUENCES: 68
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/685,871
 ; FILING DATE: 24-JUL-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 8-184102
 ; FILING DATE: 25-JUN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-262553
 ; FILING DATE: 14-SEP-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/845
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1354 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-655-871-2

Query Match 3.9%; Score 127.5; DB 3; Length 1354;
 Best Local Similarity 22.2%; Pred. No. 0.012; Matches 88; Conservatve 52; Mismatches 191; Indels 65; Gaps 16;

Qy 7 NVRNLFEQLVRYVFLSEG---NEVOFQLAKDEFREKWWORTDHELGKYDLIMKATE 64
 Db 983 NLKAFERKINNTERFLKTQAVNKLAEINRDKDFDRKANTD-----LRKEKE 1033

Qy 65 RSADVKURHARNOVYDVEIKRORADECEKL-ERQIQLREMIMQDTSGSIOLSEOKSALAFLNRG 124
 Db 1034 NRKLQLELNQERERKFNMQVKHOR----ELNDQAOVLEE-CAHRENLOQMASKE- 1084

Qy 125 LAFLNRGQPSSNNAGNKRKLSTIDESGSLSDISFDKTDLESQDWSLUVTFKLKRE--K 182

RESULT 11
US-07-906-349A-5
Sequence 5, Application US/07906349A
; Patent No. 5434064

GENERAL INFORMATION:
 APPLICANT: Schlessinger, Joseph
 APPLICANT: Margolis, Benjamin L.
 TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AND NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brodsky and Neimark
 STREET: 419 Seventh Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004

COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07-906,349A
 FILING DATE: 30-JUN-1992
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/643, 237
 FILING DATE: 18-JAN-1991
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-628-5197
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
 LENGTH: 724 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-07-906-349A-5

Query Match 3.9%; Score 126.5; DB 1; Length 724;
 best Local Similarity 22.2%; Pred. No. 0.0054;
 Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

QY 325 PECRRCRCPICPICTLGPVKIGEMLA-DFVTSOTSP-MIPSIVYHCVNEIEQRGLTE 380
 QY 325 PECRRCRCPICPICTLGPVKIGEMLA-DFVTSOTSP-MIPSIVYHCVNEIEQRGLTE 380
 Db 87 PKPRPRPRLPVAPGSSKTEADVEQQALTLPLDIAEOFAPPDIAPPLKLVAIEKKGLEC 146
 QY 381 TGLYRISGCDRTVKELKEFLRVKTVPLSKVDDTHAIACSLKDFRLNLKEPLTFLRNR 440
 Db 147 STLYKQTS SSNLARLQ-LDCDTPSVDEIMDWHVLADAKRVLLEPKLQLSQTSSKNLNARVL 204
 QY 441 AFME-AAETIDEDNSITAAMYAV-GELPGANRDLAFMLMHLYRAQ-SFHTKMDVANL 496
 Db 87 PKPRPRPRLPVAPGSSKTEADVEQQALTLPLDIAEOFAPPDIAPPLKLVAIEKKGLEC 146
 QY 205 EMISLAPEVOSSEYIQLKKLRISPSIHPQWLTQYLKLKFLSQTSSKNLNARVL 264
 QY 497 AKVFGPTIVAHAVPNPDVPMQSDTKROPKVVERILSLPLWEWSQ 541

RESULT 12
US-08-167-035-2
Sequence 2, Application US/08167035
; Patent No. 5618691

GENERAL INFORMATION:
 APPLICANT: Schlessinger, Joseph
 APPLICANT: Skolnick, Edward Y.
 APPLICANT: Margolis, Benjamin L.

TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AND NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:
 ADDRESSEE: PENNIE & EDMONDS
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: 10036-2711
 ZIP: 10036-2711

COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/167,035
 FILING DATE: 16-DEC-1993
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30-742
 REFERENCE/DOCKET NUMBER: 7683-062

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 724 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein

US-08-035-2

Query Match 3.9%; Score 126.5; DB 1; Length 724;
 best Local Similarity 22.2%; Pred. No. 0.0054;
 Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

QY 325 PECRRCRCPICPICTLGPVKIGEMLA-DFVTSOTSP-MIPSIVYHCVNEIEQRGLTE 380
 QY 381 TGLYRISGCDRTVKELKEFLRVKTVPLSKVDDTHAIACSLKDFRLNLKEPLTFLRNR 440
 Db 147 STLYKQTS SSNLARLQ-LDCDTPSVDEIMDWHVLADAKRVLLEPKLQLSQTSSKNLNARVL 204
 QY 441 AFME-AAETIDEDNSITAAMYAV-GELPGANRDLAFMLMHLYRAQ-SFHTKMDVANL 496
 Db 87 PKPRPRPRLPVAPGSSKTEADVEQQALTLPLDIAEOFAPPDIAPPLKLVAIEKKGLEC 146
 QY 205 EMISLAPEVOSSEYIQLKKLRISPSIHPQWLTQYLKLKFLSQTSSKNLNARVL 264
 QY 497 AKVFGPTIVAHAVPNPDVPMQSDTKROPKVVERILSLPLWEWSQ 541

Db

265 SEIFSPMLFRFSAASSD----NTENLIKVIELISTE--WNE 300

RESULT 13

US-08-208-887A-2
Sequence 2, Application US/08208887A; Patent No. 5677421
; GENERAL INFORMATION:

; APPLICANT: Schlessinger, Joseph

; APPLICANT: Skolnick, Edward Y.

; APPLICANT: Margolis, Benjamin L.

; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR

; IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE

; KINASES AND NOVEL TARGET PROTEINS

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/208, 887A

; FILING DATE: 11-MAR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7683-063

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 724 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-208-887A-2

; Query Match

; Best Local Similarity 3.9%; Score 126.5; DB 1; Length 724;

; Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

; Sequence 2, Application US/08339005

; Patent No. 5858686

; GENERAL INFORMATION:

; APPLICANT: App, Harold

APPLICANT: Schlessinger, Joseph

APPLICANT: Skolnick, Edward Y.

APPLICANT: Margolis, Benjamin L.

TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR

IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE

KINASES AND NOVEL TARGET PROTEINS

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/539, 005

FILING DATE: 4-OCT-1995

CLASSIFICATION: 435

PRIMER APPLICATION DATA:

APPLICATION NUMBER: US 08/167, 035

FILING DATE: 16-DEC-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7683-062

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 724 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-539-005-2

Query Match

Best Local Similarity 3.9%; Score 126.5; DB 2; Length 724;

Matches 50; Conservative 50; Mismatches 106; Indels 19; Gaps 9;

Sequence 2, Application US/09280598

Patient No. 6391584

GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph

APPLICANT: Skolnick, Edward Y.

APPLICANT: Margolis, Benjamin L.

APPLICANT: App, Harold

RESULT 14
US-08-539-005-2
Sequence 2, Application US/08339005
; Patent No. 5858686
; GENERAL INFORMATION:RESULT 14
US-08-539-005-2
Sequence 2, Application US/08339005
; Patent No. 5858686
; GENERAL INFORMATION:

TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AND NOVEL TARGET PROTEINS
TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:

ADDRESSEE: *Felicity & Edmund*, Americas
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, version #1 255

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

FILING DATE: 02-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.

REFERENCE/DOCKET NUMBER: 7683-067
TELECOMMUNICATION INFORMATION:
PHONE: (212) 790-9090
FAX: (212) 790-0741
TELETYPE: (212) 790-9090

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

Query Match	3.9%	Score 126.5;	DB 4;	Length 724;
Best Local Similarity	22.2%	Pred. No. 0.0054;		
Matches 50;	Conservative 50;	Mismatches 106;	Indels 19;	Gaps 9;
Qy	325	PECRDRCDLPCIPCLTLIGTPVKLGEGMLA--DFVSQTS--MIPSIVVHCNEIQRGLTE	380	
Db	87	PRPRPRPRLPVARGSKTEADYDQOALITPLDAEQFAPPDIAPPLLIKLYVEALEKGLEC	146	
Qy	381	TGLYRISGCDRTVKELKEFLRVKTVPLSKVDDIHATCSLLKDFLNLEKEPLTFRNR	440	
Db	147	STLYRTOS-SSNLAELRO-LLDCDTPSVDEMLIDVHVILADAFKRYLLDLPNPVPAAYS	204	
Qy	441	AEME-AEITDEEDNSIAAMQAV-GELPQANRDTLAFMLIMHQRAVA-SPHTKDVNL	496	
Db	205	EMISLADEVQSSYEIQJQLKLIRSPSAPHQYMIWLTQQLKKHFKLSQTSSKNLNLARVL	541	
Qy	497	AKVFGPPTIVAHAVPNPDFTVMSODIKRKPQKVVERLULSIPLEYNSQ	541	
Db	265	SELFSPMLFRFSAASTD----NTENIJKVINITLSTE--WNE	300	

Search completed: July 3, 2003, 09:56:51
Job time : 31 secs